Desh

Vol. 47 MAY, 1956 No. 5

The Journal of Educational Psychology

UNIVERSITY OF MICHIGAN
DEC 20 1956
EDUCATIONAL READING ROOM

CONTENTS

Some Effects of Short Term Training in Reading under Conditions of Controlled Motivation 257
JAMES C. REED
A Note on Factor Models 265
CHESTER W. HARRIS
The Davis-Eells Test and Reading Success in First Grade 269
IVAN L. RUSSELL
The Minnesota Teacher Attitude Inventory in Evaluating the Teaching of Educational Psychology 271
MORRIS E. ESON
Recency and Type of Undergraduate Training and Decline of G. R. E. Performance with Age 276
R. TRAVIS OSBORNE AND WILLIAM B. SANDERS
The Development of New Scales for the Aptitude and Advanced Tests of the Graduate Record Examinations 285
MARGARET K. SCHULTZ AND WILLIAM H. ANGOFF
The Learning of Subject Matter in Teacher-Centered and Group-Centered Classes 295
GERARD V. HAIGH AND WARREN SCHMIDT
A Study of the Family Background of the Gifted 302
WALTER B. BARBE
Determining the Relationship between Individual Interest Profiles and Occupational Forms 310
ALBERT J. KINGSTON, CLAY E. GEORGE AND W. PRICE EWENS
Book Reviews 317
Publications Received 320

Published monthly except June to September
$7.00 a year in U. S. and Pan America; Canada, $7.20; other countries, $7.40.
Single issues $1.20

WARWICK & YORK, INC.
BALTIMORE 2, MD.

Entered as Second Class Matter Nov. 15, 1921, at the Post Office at Baltimore, Md. under the Act of March 3, 1879.

THE JOURNAL OF Educational Psychology

Established 1910

BOARD OF EDITORS*

WM. CLARK TROW, *Chairman*
University of Michigan

STEPHEN M. COREY
Teachers College
Columbia University

J. B. STROUD
State University of Iowa

*Appointed by the American Psychological Association.

EDITORIAL CONSULTANTS

PAUL BLOMMERS
CHESTER W. HARRIS
HAROLD E. JONES
H. H. REMMERS
PERCIVAL M. SYMONDS
MILES A. TINKER
ALEXANDER G. WESMAN
PAUL A. WITTY

THE JOURNAL OF EDUCATIONAL PSYCHOLOGY is devoted primarily to the scientific study of problems of learning, teaching, and measurement of the psychological development of the individual. THE JOURNAL will contain articles on the following subjects: the psychology of school subjects; experimental studies of learning; the development of interests, attitudes, and personality, particularly as related to school adjustment; emotion, motivation, and character; mental development and methods. This last will include tests, statistical techniques, and research techniques in cross-sectional and developmental studies.

INFORMATION FOR CONTRIBUTORS

Manuscripts and communications regarding editorial matters should be addressed to the Chairman of the Board of Editors, Wm. Clark Trow, School of Education, University of Michigan, Ann Arbor, Mich.

THE JOURNAL has set regulations regarding content and style of material published, and these should be observed in the preparation of manuscripts to be submitted.

Tables and graphs.—In keeping with the practice of most professional periodicals of like character, THE JOURNAL now requires authors to bear part of the added cost resulting from the use of tables, formulas, and graphs. Prospective contributors may obtain detailed statement as to

(Continued on Inside Back Cover)

THE JOURNAL OF

EDUCATIONAL PSYCHOLOGY

Volume 47 May, 1956 Number 5

SOME EFFECTS OF SHORT TERM TRAINING IN READING UNDER CONDITIONS OF CONTROLLED MOTIVATION

JAMES C. REED[1]

University of Chicago

In recent years a number of attempts have been made to evaluate reading improvement training given to college students. By and large, the results of these studies have agreed in showing significant increases in rate. Some writers report a gain beyond chance in comprehension as measured by standardized tests. A few studies have indicated that such training has yielded measurable change in over-all academic proficiency as reflected by an improvement in grade point average. These results are encouraging, and at times provide information concerning the nature of the reading process. However, too often it is not at all clear that the reported gains were due to the interpolated remediation. Errors in methodological design and misuse of statistical techniques have imposed serious limitations on the generalizations that can be drawn from many of the studies as well as their validity. The two most serious weaknesses have been, first, a failure to equate the experimental and control groups on the basis of initial motivation, and secondly, a failure to randomly divide the participants. The difficulties encountered have been that volunteers have been paired with non-volunteers thereby permitting a bias. Also, the statistics t and F have been used to test the gains, and these values are based on statistical theory which assumes randomly chosen observations.

McGinnis' (*8*) experimental cases were volunteers to the reading clinic. Her control group was selected from the college at large, and though the two were carefully matched on the basis of sex, status in college, scholastic aptitude, and level of reading, they were not

[1] On leave of absence from Wayne University.

equated for initial motivation and the principle of random selection was violated. She found that the trainees exceeded the untrained in grade point average.

Kilby (*6*) also reported that instruction in reading resulted in significantly higher grades. He attempted to control for motivation by including among his matching criteria a "predicted grade score." It is questionable whether a predicted grade score derived from College Board Entrance Examinations, an achievement test, and adjusted high school grade-point is an adequate measure of motivation, and its applicability as such is suspect since Kilby's trained group were apparently volunteers and the controls picked from the general college population.

In Barbe's study (*2*), only those who elected reading were used. His experimental group consisted of the first twenty-five volunteers who were able to include the course in their schedules. The controls were the first twenty-five who could not so modify their program. While he reported significant gains in speed of reading and in honor point average for the experimental subjects, there is no evidence to indicate that the two groups were equivalent in intellectual ability. In addition, the method Barbe used for computing the significance of the rate gains will not yield a precise measure of the net shift in rate of reading. Furthermore, the possibility of differential motivation exists and the subjects were not randomly assigned. Robinson (*11*) likewise used only volunteers. However, the design was better in that control group which consisted of those who agreed to take the course at a later date. Unfortunately more than twenty-five per cent of them did not return for the class. Though Robinson's two groups were not based on random assignment, there is reason to believe that the factor of motivation was more nearly equated than in the studies heretofore mentioned. It is interesting to note that the grade point differences were not statistically significant although they approached being so. Mouly (*10*) saw the necessity of randomly sorting his subjects into experimental and control groups. He encountered difficulty in keeping motivation equal. Some of the students required to take the reading training refused to do so. Moreover, he found significant difference in grade point only when the experimental group was limited to those members who successfully completed the course. When the drop-outs and failures were included, the differences were no longer significant.

Such procedures as have been described suggest that the reported gains may be a function of characteristics of the participants rather than of treatments. The magnitude of gain seems inversely related to the degree of precision with which the experimental and control groups were matched. It may be that the traits which lead undergraduates to seek reading instruction are the same ones that result in higher grades than their less ambitious peers.

In the present experiment the hypothesis to be tested was that under conditions of controlled motivation, intensive training in reading and study skills would yield significant gains in rate, vocabulary, comprehension, and honor point average.

PROCEDURE

The subjects used were taken from the entering class of freshman hospital contract nursing students at Wayne University. From data contained in the entrance examination, the cases were matched by pairs on the basis of their ACE Q score (*1*) and their performance on the vocabulary and comprehension sections of the Diagnostic Survey Test, Form A (*3*). It was possible to find thirty-six individuals, eighteen pairs, whose intra-pair difference was no more than 5 raw score points on each of the three variables. The means for these thirty-six on ACE Q score, on comprehension, and on vocabulary were respectively at the twenty-fifth, fifty-first, and fiftieth percentile ranks on Wayne University freshman norms. By randomly assigning each member of each of the eighteen pairs into an experimental and control group, the differential motivation associated with the use of volunteer subjects was precluded. To determine whether there was comparability on the matching variables, means and standard deviations were computed. The difference between the means of the two groups on each of the three factors did not exceed 0.6 of a point. The sigmas were likewise identical.

After the experimental and control subjects had been selected, the entire class of freshman nurses was administered the Untimed Comprehension Test (*3a*) and the Vocabulary Test from the Diagnostic Battery (*3b*). On neither of these measures did the mean difference between the experimentals and controls exceed chance. It was announced that the course in reading improvement was required for student nurses, that there would be two sections—one

now and the other at a later date—and that their placement into a section would be purely by chance.

The reading class was composed of twenty-five members, eighteen of whom were the experimental subjects. The other seven were randomly chosen from those for whom no pair could be found at the time of the original matching. The training consisted of practice in rate of reading, drill on comprehension exercises, instruction in vocabulary, and discussion of study techniques. Glock's (5) *The Improvement of College Reading* was used as a text for the course supplemented by a wide variety of other material. The group met for an hour three times a week for nine weeks—a total of twenty-seven hours.

At the close of the training, alternate forms of the original tests were administered along with the Survey Section, Form B. A week later the control group and the remaining unpaired portion of the nursing class were separately given the same post-tests. The examiner stated that the second session of the course was to be offered during the spring semester and that at this time he wanted to determine how much gain was made in reading by students when

TABLE I.—DIFFERENCES IN GAIN (G) MADE BY EXPERIMENTAL AND CONTROL GROUPS

$N = 18$

Tests	Experimental					Control					Differences		
	Beginning		Ending			Beginning		Ending					
	M	σ	*M*	σ	G_e	*M*	σ	*M*	σ	G_c	$D_{G(e-c)}$	σdiff.	*t*
Survey section													
Rate	268	33.25	353	59.83	85	249	39.63	285	42.69	36	49	20.20	2.44*
Vocab.	43	6.76	48	6.83	5	43	6.71	39	7.80	−4	9	2.46	3.61†
Comp.	30	4.39	30	2.85	0	30	5.30	31	3.42	1	−1	1.02	<1
Sec. 1 Vocab.	110	12.02	108	13.76	−2	116	7.99	109	9.76	−7	5	2.75	1.80
Sec. 2 Comp.	35	5.80	34	4.92	−1	37	5.20	35	4.78	−2	1	1.51	<1
Honor point			2.39	0.79				2.49	0.68		−0.1	0.19	<1

* Significant at two per cent.

† Significant beyond one per cent.

they did not receive special training, because if the gains were great enough, it was possible that the course would not be required.

The net gains in rate, vocabulary, and comprehension were calculated for each subject. The differences in net gain within each pair of subjects were computed for each of the three dependent variables. Using these distributions of differences, *t* tests were run. At the conclusion of the fall semester, the honor point averages of the subjects were obtained, and the differences were similarly evaluated. (*9*, p. 78).

Seven months after the end of the training period, the entire group was reassembled, and administered Form C of the Survey Section, and the same forms of the other tests that had been given to them initially. The statistical analysis of these delayed post-test results was done on the distribution of intra-pair difference scores of the experimental and control subjects.

RESULTS

As is contained in Table I, on the Survey Section of the Diagnostic Reading Tests, the experimental subjects made statistically significant gains over their paired controls in rate of reading, 0.02 per cent level, and on the vocabulary part, beyond the 0.01 per cent level. There were no significant differences found between the experimental and control groups on the vocabulary tests from Section 1 of the Diagnostic Battery, in comprehension, and in honor point average. In addition to the *t*, a nonparametric sign test as described by Dixon and Massey (*4*) was done for the pairs on each of the dependent measures. For rate of reading, the sign test was significant beyond the 0.01 confidence level. The remaining signs were not significant.

From Table II it can be seen that at the end of seven months the experimental group maintained its superiority in rate. The significant *t* between the two groups in the vocabulary section of the Survey Test was no longer present and none of the remaining differences exceeded chance. As before, only the sign test for rate of reading scores was significant.

CONCLUSIONS

The manner by which the subjects were selected for this experiment minimized the possibility that the experimental group initially had a higher degree of interest in the course or were more

TABLE II.—NET DIFFERENCES (D) BETWEEN EXPERIMENTAL AND CONTROL GROUPS SEVEN MONTHS AFTER TRAINING

$N = 14$

Tests	Experimental		Control		Differences		
	M_e	σ	M_c	σ	$D_{M(e-c)}$	σdiff.	t
Survey section							
Rate	320	33.25	283	22.66	37	9.05	4.07*
Vocab.	40	4.34	40	4.47	0	1.61	<1
Comp.	30	4.53	31	3.05	−1	1.46	<1
Sec. 1 Vocab.	116	9.73	119	7.46	3	2.48	1.10
Sec. 2 Comp.	35	5.97	37	5.84	−1.5	2.20	<1

* Significant beyond one per cent.

strongly motivated than the control group as frequently happens when volunteers are employed. Judging from the response of the nurses in the classes, their number of absences, the frequency of late comers, their postural indications of attention, it is believed that the group interest and motivation was equal to that of an average introductory "required" course. The nurses were somewhat inclined to be phlegmatic and they did what they were told without too much question. Consequently, it is felt that the changes in reading skills are due to the training and not to antecedent conditions peculiar to the experimental subjects.

The conclusions and generalizations that may be drawn from these results are necessarily limited due to the small size of the N. It should be remembered that the sample is not representative of a college population. However, it would appear that the following conclusions are justified.

(1) The significant increase in rate of reading without change in comprehension or vocabulary suggests that rate is an independent factor in the reading process. Furthermore, rate will increase as a result of training and this increase has some degree of permanence.

(2) Since a significant difference was found on the timed vocabulary test of the Survey Section but not on the more liberally timed ones of Section 1, it appears that the vocabulary section of the Diagnostic Survey Test is in part a measure of rate. Such an interpretation is further supported by the fact that the difference on the Survey vocabulary test between the experimental and controls was not significant on the delayed post-tests.

(3) Under appropriate matching and equating of initial motivation, short-term training in reading does not yield material differences in fundamental comprehension and vocabulary skills.

(4) It would likewise appear that changes in honor point averages are not brought about by reading training. Grades are a function of a complex interaction of many variables, and are in part due to habits and abilities of long standing. It would be rather surprising if as the result of a nine-weeks course, there were substantial differences in academic achievement.

(5) It may be that failure to achieve significant changes in grade point is due to the specialized curriculum in which the nurses were enrolled. Their courses were mainly in the biological and physical sciences where quite intensive and comprehensive reading is required. It is possible that the training as described herein would be beneficial in terms of an increase in honor point for students enrolled in the social sciences and humanities. Credence is given such an interpretation by the works of Sheldon (*12*), Kilby (*6*), and Kingston (*7*). However, the writer does not share the view that grade points are a sole or necessarily even a useful criterion in evaluating a reading program.

(6) It is recognized that discrepancies between the results of this and other experiments may possibly be due to differences in instructional procedures, the length of time the classes were offered, the type of students, and even the competence of the instructors. Yet, it is felt that this experiment has demonstrated a need for more careful evaluations of reading programs at the college level. More attention should be directed to the appropriate matching of experimental and control subjects. The precision of the experiments can be increased by the use of appropriate estimates of error. Finally, repetition of such studies must be done in order to randomize the types of error not possible to eliminate or randomize by any single experiment.

REFERENCES

1) American Council on Education, *Psychological Examination*. Cooperative Test Division, Educational Testing Service, Princeton, New Jersey.

2) Walter H. Barbe, "The Effectiveness of Work in Remedial Reading at the College Level," *Journal of Educational Psychology*, 43: 229–237, 1952.

3) *Diagnostic Reading Tests*. Survey Section. Committee on Diagnostic Reading Tests, Kingscott Apartment, 3 G, 419 West 119th Street, New York 27, New York.

3a) *Diagnostic Reading Tests*, Section II, Comprehension, Part I Silent.

3b) *Diagnostic Reading Tests*, Section I, Vocabulary.

4) Wilfred J. Dixon and Frank J. Massey, Jr., *Introduction to Statistical Analysis*, pp. 247–254. New York, McGraw-Hill, 1951.

5) Marvin D. Glock, *The Improvement of College Reading*. Chicago, Houghton-Mifflin Co., 1954.

6) R. W. Kilby, "The Relation of a Remedial Reading Program to Scholastic Success in College," *Journal of Educational Psychology* 36: 513–534, 1945.

7) Albert J. Kingston and George E. Clay, "The Effectiveness of Reading Training at the College Level," *Journal of Educational Research* 48: 467–471, 1955.

8) Dorothy McGinnis, "Corrective Reading: A Means of Increasing Scholastic Attainment at the College Level," *Journal of Educational Psychology*, 42: 166–173, 1951.

9) Quinn McNemar, *Psychological Statistics*, p. 78. New York, J. Wiley, 1949.

10) George J. Mouly, "A Study of the Effects of a Remedial Reading Program on Academic Grades at the College Level," *Journal of Educational Psychology*, 43: 459–466, 1952.

11) H. A. Robinson, "A Note on the Evaluation of College Remedial Reading Courses," *Journal of Educational Psychology*, 41: 83–96, 1950.

12) William Sheldon, "An Evaluation of an Experimental Reading Program for Medical Students," *Journal of Educational Psychology* 39: 298–303, 1948.

A NOTE ON FACTOR MODELS

CHESTER W. HARRIS

University of Wisconsin

It is important to distinguish, as Maxwell (*9*) has done, between the choice of a factor model and the choice of computation procedure. His preference, like that of many others (*2*, *3*, *4*, *7*, *11*), is for the communality principle. The purpose of this note is to emphasize two points: (1) that two well-known systems of factor analysis—i.e., factor models—though distinguishable, are related in a regular fashion, and (2) that computation procedures which are readily performed on digital computers, such as those described by Wrigley and Neuhaus (*12*), may be utilized to secure a solution utilizing the communality principle.

FACTOR MODELS

The distinction between two systems of factor analysis, or factor models, can be visualized by considering the trivial problem of only two variables. One approach is that of factoring the two-by-two matrix consisting of units in the principal diagonal and the correlation between the two variables in the off-diagonal positions. This takes as a complete initial orthogonal solution a first factor midway between the two test vectors but lying in the plane of these two vectors, and a second factor at right angles to this first factor but again lying in the plane of these two test vectors. This approach locates the factors *within the test space*. A second approach is given by the communality principle, which corresponds to factoring correlations with communalities in the principal diagonal. Three factors are identified, two of which are unique to the two variables. The single common factor is located at the intersection of two orthogonal planes, each of which passes through one of the test vectors, and consequently is located *outside the test space*. It also is true that the unique factors are located outside the test space. These two approaches, then, differ in that for one the factor space is defined as identical with the test space, whereas for the other the common-factor space is defined as distinct from the test space.

A number of qualifying comments should be introduced at this

point. For example, it should be remarked that the two approaches give identical results under the condition of perfect correlation between the two variables, and the effects of using covariances rather than correlations should be considered.

It has been shown (*6*) that this distinction in factor models corresponds to the distinction between factoring the given data, such as a normalized set of test scores for N individuals, and factoring a projection of this matrix of data. Both procedures imply a separation of the given data into two parts, which is the basic statistical operation in analysis of variance and in linear regression. The communality principle in factor analysis can therefore be formulated as a straight-forward statistical problem. Anderson and Rubin (*1*), Lawley (*8*), and Rao (*10*) have discussed alternative ways of separating the given data into two parts and the types of solutions these imply.

Given the distinction between a set of orthogonal factors located within the test space and one or more common factors located outside the test space by using the communality principle, the interesting question of the relationship between these two distinct sets of factors arises. As an illustration, consider a two-variable matrix:

$$\begin{Vmatrix} 1.00 & 0.64 \\ 0.64 & 1.00 \end{Vmatrix}$$

The communality solution yields a single common factor, here labeled G, and the complete principal-axis (or centroid in this case) solution yields two common factors, here labeled I and II:

Variable	G	I	II
1	0.8	$\sqrt{0.82}$	$\sqrt{0.18}$
2	0.8	$\sqrt{0.82}$	$-\sqrt{0.18}$

The transformation that projects the complete set of factors, I and II, into the single common factor, G, can readily be computed for this simple problem. This transformation is simply a single row vector with a non-zero entry, consisting of $\sqrt{0.64}$ divided by $\sqrt{0.82}$, in the first cell and a zero in the second cell. More generally, it can be shown (*5*) that for population values of the correlations for which a communality solution exists, such a trans-

formation always exists and consists only of real numbers. When this relationship is considered in connection with the problem of estimating factors, it turns out that the transformation can be defined algebraically, and that common factors solved for by this transformation satisfy Lawley's maximum likelihood requirement and are a first approximation to Rao's canonical factors.

This relationship between two systems of factor analysis therefore, in one sense, blurs the distinction in factor models emphasized by Maxwell, since it shows the mode of transforming one to the other; yet at the same time it helps sharpen the distinction by specifying the projective nature of this transformation.

COMPUTATION PROCEDURES

Implicit in this discussion is the notion that one might factor completely any given correlation matrix with units in the diagonal cells, and then by further calculation determine common factors as a transformation of this complete set of factors. (Variances and covariances might be used instead, and the resulting factors would be simply related to those determined by operating on the correlations.) This might be done by machine, especially since all of the factoring required by this process should be done by the principal-axis transformation, which is well-adapted to digital computing. However, this process actually is uneconomic compared with the direct attack on the estimation of common factors given by Rao (*10*) in his notion of canonical factors. His computation procedure also requires the principal-axis transformation of a matrix, and thus is well-adapted to digital computing.

Perhaps the best advice for the factor analyst today is, "Study Rao's canonical factor analysis procedure and whenever the matrix to be factored is of any size, arrange for machine computation, employing Rao's method." Operations appropriate to high-speed computing machines are employed, a "respectable" mathematical model is used, and appropriate tests of statistical significance are made available. In addition—and this is not insignificant—the communality principle is preserved.

REFERENCES

1) T. W. Anderson and Herman Rubin, *Statistical Inference in Factor Analysis*. Technical Report No. 30, Contract N6onr-251, Office of Naval Research. Stanford, California, Department of Statistics, Stanford University 1955.

2) Raymond B. Cattell, *Factor Analysis*. New York, Harper and Brothers, 1952.

3) H. J. Eysenck, "The Logical Basis of Factor Analysis," *The American Psychologist*, pp. 105–114, March, 1953.

4) J. P. Guilford, *Psychometric Methods*. New York, McGraw-Hill, 1954.

5) Chester W. Harris, "Relationships Between Two Systems of Factor Analysis," *Psychometrika*, (in press).

6) Chester W. Harris, "Separation of Data as a Principle in Factor Analysis," *Psychometrika*, pp. 23–28, March, 1955.

7) Karl J. Holzinger and Harry S. Harman, *Factor Analysis*. Chicago, University of Chicago Press, 1941.

8) D. N. Lawley, "The Estimation of Factor Loadings by the Method of Maximum Likelihood," *Proceedings, Royal Society of Edinburgh* pp. 64–82, 1940.

9) A. E. Maxwell, "Factor Models," *Journal of Educational Psychology*, p. 9 March, 1956.

10) C. R. Rao, "Estimation and Tests of Significance in Factor Analysis," *Psychometrika*, June, 1955, 93–111.

11) L. L. Thurstone, *Multiple-Factor Analysis*. Chicago, University of Chicago Press, 1947.

12) Charles Wrigley and J. O. Neuhaus, "Use of an Electronic Computer in Principal Axes Factor Analysis," *Journal of Educational Psychology*, pp. 31–41, January, 1955.

THE DAVIS-EELLS TEST AND READING SUCCESS IN FIRST GRADE

IVAN L. RUSSELL

Southern Illinois University

The Davis-Eells Test is a unique approach to the measurement of intelligence. Attention is focused upon the child's capacity for solving problems similar to true-life experiences. An attempt is made to minimize the influence of socio-economic differences. The purpose of the Davis-Eells Test is to "afford the teacher a measure of a pupil's mental capacity, of his basic resources of reasoning, insight, problem organization and thus to throw new light for her upon the pupil's true capacity." (*1*)

A study was conducted to test the hypothesis that the Davis-Eells Test provides important information regarding reading readiness. The plan of study included administration of the Davis-Eells Test to a randomly selected group of fifty first-grade pupils during the first week of school. The test was administered to groups of ten pupils each, and the directions were followed carefully. For comparative purposes the Stanford-Binet Test, Form L, was administered to each of the fifty pupils during the first two weeks of school. During the last two weeks of school in May the Gates Sentence and Paragraph Reading Tests were given. These sets of test scores were compared statistically with the Gates reading scores, which were obtained at the end of the first grade, as the criterion of reading success.

PRESENTATION OF DATA

Table I compares the means and standard deviations of IQ's earned on the Stanford-Binet and Davis-Eells Tests.

These differences between the means on the two tests are significant at the one per cent level. The difference between mean Davis-Eells scores for boys and for girls is significant at the five per cent level.

A Pearson product-moment correlation coefficient of + 0.48 was obtained between the Davis-Eells IPSA (Index of Problem Solving Ability) and Stanford-Binet IQ scores. Stanford-Binet mental ages correlated with Gates reading ages obtained at the

TABLE I.—MEANS AND STANDARD DEVIATIONS FOR DAVIS-EELLS AND STANFORD-BINET SCORES

Group	*N*	Davis-Eells (IPSA)		Stanford-Binet (IQ)	
		Mean	S.D.	Mean	S.D.
Girls	28	95.3	12.2	110.0	14.5
Boys	22	87.9	12.4	104.5	14.6
Total	50	91.9	13.1	108.9	14.8

end of the school year are as follows: for boys, + 0.48; for girls, + 0.63; for total group, + 0.57.

The Davis-Eells test does not provide M.A. scores. These were derived by multiplying each pupils IPSA score, roughly equivalent to IQ, by his chronological age. The M.A.'s thus derived were correlated with end-of-the-year reading scores, just as in the case of the Stanford-Binet Test. The coefficients of correlation which resulted were: girls $r = 0.28$; boys $r = 0.08$; total group $r = 0.21$.

CONCLUSIONS

(1) IPSA scores obtained in this study involving first-grade pupils proved to be significantly lower than Stanford-Binet IQ's ($p = 0.01$).

(2) For this limited sample the Stanford-Binet test gives a better prediction of reading progress during the first year's instruction than the Davis-Eells Test.

REFERENCES

1) Allison Davis and Kenneth Eells, *Manual: Davis-Eells Test of General Intelligence or Problem-Solving Ability*, p. 2. New York, World Book Co., 1953.

THE MINNESOTA TEACHER ATTITUDE INVENTORY IN EVALUATING THE TEACHING OF EDUCATIONAL PSYCHOLOGY

MORRIS E. ESON[1]

New York State College for Teachers, Albany, N. Y.

One of the important objectives of the professional courses in teacher education is the development and change of attitudes toward the learner and the teaching task. The teacher of such courses may be reluctant to measure the extent to which this objective is achieved because there have been no well-standardized instruments for the measurement of teacher attitudes and also because the connection between the content of the professional courses and any set of attitudes is not always clear. The study reported here represents an attempt to measure the attitude change which took place in one such course, the first in the professional sequence at New York State College for Teachers, Albany, N. Y. This course, Education 20, is generally taken by the student during the first semester of the sophomore year and follows a three-semester-hour course in General Psychology. It is a course in Educational Psychology and is described in the 1954–55 college catalogue as follows:

> The study of basic concepts of human behavior as related to the educational situation. Special emphasis will be given to physical, social, emotional and mental development in early and later childhood; the study of the interaction of heredity and environment in the development of personality; the learning process including a consideration of reasoning, remembering, transfer of training and the various theories of learning. One two hour period each week will be spent in laboratory session to develop insight into human behavior by observation, preparation of reports and other laboratory devices.

The text used in the course was James L. Mursell's *Psychology for Modern Education*, W. W. Norton & Co., 1952. The writer met the classes twice each week for sixteen weeks. These class sessions

[1] The writer gratefully acknowledges the help of Owen Smith and Nancy D. Westover, graduate students at New York State College for Teachers, in the statistical analysis of the data.

were spent in lecture, discussion and demonstration. Each student observed five times in elementary school classrooms and five times in secondary school classrooms and was required to submit observation reports on seven of these observations. These reports were returned with comments by the instructor. The instructor's philosophy may be described as consistent with that which would be scored high on the Minnesota Teacher Attitude Inventory. The instructor's score on the Inventory was such as to place him at the ninety-fifth percentile on the norms established for secondary school teachers with five years of training.

The population used in the study consisted of seventy-five first-term sophomore men and women who were taking their first professional course in education, Educational Psychology. The data used in the study were collected as follows: a) During the first class session the Inventory was administered. b) After eight weeks an instructor-constructed objective test covering the material in approximately the first half of the text was administered. c) The Inventory was again given during the last class session. d) A final objective test, instructor constructed, was given in the final examination week. At each of the administrations of the Inventory, the students were assured that their scores would in no way influence their final grade. The response sheets on the Inventory were machine scored.

The rationale of the Inventory may best be described by a series of quotations from the Manual (*1*).

It is assumed that a teacher ranking at the high end of the Scale should be able to maintain a state of harmonious relations with his pupils characterized by mutual affection and sympathetic understanding.

The teacher should like the children and enjoy teaching.

The teacher and pupils should work together—with a feeling of security growing from a permissive atmosphere of freedom to think, act and speak one's mind with mutual respect for the feelings, rights and abilities of others.

At the other extreme of the Scale is the teacher who attempts to dominate the classroom.

Creating an atmosphere of tension, fear and submission.

There is a feeling of mutual distrust and hostility.

The teacher tends to think in terms of . . . the subject matter to be covered rather than in terms of what the pupil needs, feels, knows and can do.

He frequently has a submissive uncritical attitude toward authorities over him and a dominating, overbearing attitude toward subordinates.

The authors of the Inventory state that the items in the Inventory "discriminate sharply between teachers who have and those who do not have good rapport with pupils; examination of these items indicates that inferior teachers are essentially insecure socially." The factors suggested causing social insecurity are general appearance, failure in heterosexual adjustment, low social status of family, failure to be accepted socially in high school and the like. In teaching the course here being evaluated, no attempt was made to change these factors; obviously some of these could not be changed. In fact it would seem that the attitudes measured by the Inventory are resistant to change and that "the mere inculcation of better attitudes by instruction may not produce any change in behavior." It is to be speculated as to whether instruction in the principles of child development and the learning process may produce changes in attitudes as measured by the Inventory and, if so, whether these attitude changes would reflect also a change in behavior.

RESULTS

The mean score on the Inventory increased by 11.5 points, as seen in Table I, between the first and the second testing. The product moment correlation between the two sets of scores of the Inventory was found to be 0.69. The significance of the difference for correlated means was tested by obtaining the Standard Error which was found to be 2.58 giving a t value of 4.4, significant beyond the 0.01 level.

A correlation of +0.29 was found between the scores on the final content test and the second attitude inventory. This low

TABLE I—COMPARISON OF SCORES ON THE MINNESOTA TEACHER ATTITUDE INVENTORY TAKEN SIXTEEN WEEKS APART

	N	M	r	SD	Difference between Means	SE diff	t
MTAI first administration	75	41.4	0.69	26	11.5	2.58	4.4
MTAI second administration	75	52.9		28			

correlation between the course content and the attitudes measured on the inventory is further confirmed when the twenty-six students with the highest inventory percentile ranks are compared with the twenty-six lowest inventory scores; using letter grades obtained by these students in the course a contingency coefficient of +0.34 was obtained.

The mean score of the population when compared with beginning education juniors, secondary academic was found to be at about the fortieth percentile on the first testing and about the fifty-fifth percentile on the second testing of the Inventory.

Seventy-two per cent of the students in the study favorably increased their scores by more than five points whereas twenty per cent decreased their scores by more than five points.

DISCUSSION

Since the difference between the means on the Inventory was found to be statistically significant, one might be tempted to conclude that the course contributed to the improvement of attitudes. However, one might also conclude that the students had simply become test wise. The authors of the Inventory report two administrations of the Inventory with seven to ten days between testings and with standard instructions each time. The difference between means in that instance is found to be significant beyond the one per cent level of confidence. Our situation here parallels even more closely the sequence described by the authors in the Manual in which the subjects were encouraged to "fake good" on the second administration which followed the first by a four- to six-week interval. The authors report a t value of 4.13 which corresponds very closely to the 4.4 reported above.

The fact that correlation between the results on course content examinations and the Inventory scores is found to be low is further evidence that the difference between the Inventory mean scores, although statistically significant, is in reality not significant.

SUMMARY AND CONCLUSIONS

The Minnesota Teacher Attitude Inventory was administered to a group of seventy-five college sophomores at the beginning and at the conclusion of a course in Educational Psychology. The mean score was raised 11.5 points which is shown to be a statistically significant difference. The low correlation between scores on sub-

ject matter mastery and the attitude inventory as well as the possibility of faking on the attitude inventory cause the author to question the applicability of the Minnesota Teacher Attitude Inventory to the evaluation of courses which attempt to create and develop certain kinds of teacher attitudes.

REFERENCE

1) W. W. Cook, C. H. Leeds and R. Callis, *Minnesota Teacher Attitude Inventory Manual*. The Psychological Corporation, New York, N. Y.

RECENCY AND TYPE OF UNDERGRADUATE TRAINING AND DECLINE OF G.R.E. PERFORMANCE WITH AGE[1]

R. TRAVIS OSBORNE and WILMA B. SANDERS

INTRODUCTION

Previous papers (*1*, *2*, *3*) by the writers have shown that there is a significant decline in mean Graduate Record Examination scores with increased age. In a study of one thousand, eight hundred and seven graduate students, the rate of loss varied both with age and subject-matter area. Over the nineteen to sixty-five age range the science profile age-score curves showed losses ranging from zero to seventeen per cent for men and from seven to twenty-five per cent for women. Achievement scores in literature held up well for both sexes over the entire age range. The question was raised "Are there factors operating other than age or in combination with age to produce these changes?" For example, recency and type of undergraduate training might affect the rate of profile score change with age. Students who earned baccalaureate degrees in the sciences not only might be superior initially in their major fields but also might show a less marked rate of achievement loss with age than non-science majors. It is also hypothesized that, regardless of age, recent graduates will test higher than members of earlier classes. The purpose of this paper, which is subject to the limitations inherent in a cross sectional study, is to present research evidence relevant to these points.

POPULATION

The one thousand, eight hundred and seven graduate students who were in attendance at the University of Georgia between February, 1946, and February, 1952, were subjects for the study. This group included the entire graduate student body except several who were not seeking advanced degrees. The seven professional schools and the five divisions of the College of Arts and Sciences were represented. Men outnumbered women one thousand one hundred and seventy-three to six hundred and thirty-four. The

[1] Based upon a paper presented at the meeting of the American Psychological Association, New York, September, 1954.

age range of subjects was from nineteen to sixty-five with a mean of 31.76 years.

RESULTS

In Table I, Graduate Record Examination profile scores are shown fractionated by time interval between the student's baccalaureate degree and his G.R.E., disregarding age at the time of the test. Thus it is possible in Table I to compare achievement profiles of recent graduates with those of students who received their A.B.'s four, eight, or up to twenty years previously. The losses in scores found when the G.R.E. results were fractionated by age (*3*) are attenuated when the data are broken down by recency of baccalaureate degree. Only on three profiles are significant differences noted in favor of the most recent graduates. Biology, chemistry, and physics scores are elevated for students who had received their degrees within three years of the time of testing. Four years after graduation, all inter-group differences on the chemistry and physics profiles as well as most differences in biology are insignificant (Table I).

Recency of undergraduate training is not a significant factor at any age level in the areas of mathematics, social science, fine arts, and literature. In fact on the literature profile the "old grads," class of 1923 and before, had significantly superior test results compared to the four other age groups (Table I).

To test the hypothesis that type of undergraduate training affects the rate of profile score change with age, Table II was prepared. It shows the means, standard deviations, and the number of science and non-science majors at each age level. B.S. students were found to be superior on the chemistry, physics, and biology profiles at practically every age level whereas non-science students excelled in social science and the humanities at all age levels. Declines appear to be sharpest for science majors on the science profiles. Science achievement loss for non-science majors is gradual with the means for the two curriculum groups becoming more nearly equal after age thirty-nine.

In order to establish an over-all test of significance these data were further subjected to analysis of variance with proper adjustment being made for disproportionality (*4*). The hypotheses to be tested by this design are: (1) The effects of student age on the profile scores of the G.R.E. are insignificant. (2) Variations in

TABLE I—RECENCY OF BACCALAUREATE DEGREE IN RELATION TO GRADUATE RECORD EXAMINATION PERFORMANCE

Years Between Baccalaureate Degree and G.R.E.	No.	Mean St.Sc.*	Years					Years Between Baccalaureate Degree and G.R.E.	No.	Mean St.Sc.*	Years				
			0–3	4–7	8–11	12–15	16–19				0–3	4–7	8–11	12–15	16–19
Mathematics								*Social science*							
0– 3 Years	638	50.59						0– 3 Years	842	49.88					
4– 7 Years	274	49.62	NS					4– 7 Years	318	49.67	NS				
8–11 Years	169	49.30	NS	NS				8–11 Years	203	50.81	NS	NS			
12–15 Years	110	50.38	NS	NS	NS			12–15 Years	153	49.98	NS	NS	NS		
16–19 Years	76	48.23	NS	NS	NS	NS		16–19 Years	105	51.53	NS	NS	NS	NS	
20 Years and up	57	49.73	NS	NS	NS	NS	NS	20 Years and up	86	51.33	NS	NS	NS	NS	NS
Physics								*Literature*							
0– 3 Years	842	51.46						0– 3 Years	842	49.63					
4– 7 Years	318	49.64	0.01					4– 7 Years	317	50.18	NS				
8–11 Years	204	48.45	0.01	NS				8–11 Years	202	50.18	NS	NS			
12–15 Years	154	47.27	0.01	0.05	NS			12–15 Years	154	49.41	NS	NS	NS		
16–19 Years	105	48.45	0.01	NS	NS	NS		16–19 Years	105	51.71	NS	NS	NS	NS	
20 Years and up	87	48.56	0.05	NS	NS	NS	NS	20 Years and up	87	52.92	−0.01	−0.05	−0.05	−0.01	NS
Chemistry								*Fine arts*							
0– 3 Years	819	51.85						0– 3 Years	829	50.21					
4– 7 Years	274	48.00	0.01					4– 7 Years	308	50.11	NS				
8–11 Years	180	47.63	0.01	NS				8–11 Years	201	49.51	NS	NS			
12–15 Years	135	47.25	0.01	NS	NS			12–15 Years	154	49.41	NS	NS	NS		
16–19 Years	97	48.74	0.01	NS	NS	NS		16–19 Years	104	50.01	NS	NS	NS	NS	
20 Years and up	81	49.86	NS	NS	NS	NS	NS	20 Years and up	84	50.81	NS	NS	NS	NS	NS

Biological science																
0– 3 Years	842	52.75														
4– 7 Years	318	48.95	0.01													
8–11 Years	203	47.46	0.01	NS												
12–15 Years	154	45.73	0.01	0.01	NS											
16–19 Years	105	46.31	0.01	0.05	NS	NS										
20 Years and up	87	45.62	0.01	0.05	NS	NS	NS									

* Standard scores based on University of Georgia norms.

TABLE II—TYPE OF BACCALAUREATE DEGREE IN RELATION TO GRADUATE RECORD EXAMINATION PERFORMANCE

Age group	Physics						Biological Sciences					
	Baccalaureate degree in science			Baccalaureate degree non-science			Baccalaureate degree in science			Baccalaureate degree non-science		
	No.	Mean St. Sc.*	S.D.	No.	Mean St. Sc.*	S.D.	No.	Mean St. Sc.*	S.D.	No.	Mean St. Sc.*	S.D.
19–22	46	51.09	10.97	82	49.67	8.87	46	58.33	9.66	82	52.41	8.28
23–26	196	54.15	9.73	211	49.62	8.01	196	56.08	9.25	211	50.67	8.33
27–30	188	53.33	10.70	159	48.83	8.98	188	54.07	9.66	159	49.75	7.99
31–34	91	51.93	9.62	112	49.28	8.87	91	51.17	9.43	112	47.96	9.60
35–38	67	49.99	10.32	131	49.40	9.73	67	46.53	8.97	130	47.11	10.11
39–42	50	50.17	8.76	97	44.62	9.41	50	47.80	8.91	97	44.27	8.45
43–46	26	45.22	8.60	62	47.79	11.72	26	44.41	10.23	62	46.61	10.92
47–50	12	50.39	9.57	39	45.84	8.66	12	47.40	8.79	39	42.36	8.16
51– up	13	38.81	5.91	38	42.89	9.62	13	37.78	8.56	38	41.42	9.66
	Chemistry						Mathematics					
19–22	44	55.55	10.99	80	48.81	10.06	33	47.23	9.95	56	50.32	8.82
23–26	192	56.22	9.32	202	47.89	9.50	156	52.12	9.84	179	50.02	10.22
27–30	176	53.50	10.12	143	47.48	8.63	153	52.00	9.68	129	49.23	10.16
31–34	83	52.50	9.63	106	48.02	8.94	63	48.74	10.97	94	49.50	10.48
35–38	57	51.28	8.20	111	46.17	9.19	49	50.59	10.59	100	51.29	11.77
39–42	49	51.35	8.39	85	45.17	7.95	37	48.49	10.91	77	47.11	9.03
43–46	24	49.02	9.63	54	46.55	9.69	16	42.41	7.58	48	49.01	11.83
47–50	11	56.68	8.32	39	46.51	9.19	8	50.13	8.28	31	46.65	10.38
51– up	12	42.28	4.04	37	44.72	8.14	9	43.64	11.61	27	49.41	10.48
	Social Sciences						Literature					
19–22	46	47.65	8.81	82	52.16	10.00	46	49.01	8.30	82	55.02	9.95
23–26	196	47.08	7.98	211	52.93	9.79	196	46.63	8.52	211	54.16	9.67
27–30	188	47.09	8.96	159	51.68	10.88	186	45.45	8.90	159	52.59	10.60
31–34	91	47.46	9.69	112	51.97	10.31	91	45.97	9.07	112	51.62	10.99
35–38	67	48.09	10.05	129	52.58	11.35	67	45.78	8.96	130	51.74	9.84
39–42	50	47.49	8.91	97	50.68	10.05	50	47.32	8.46	97	51.90	9.07
43–46	25	49.57	9.02	62	51.74	10.78	26	47.26	7.86	62	54.02	11.04
47–50	12	46.54	7.46	39	47.97	8.29	12	44.87	7.03	39	48.88	8.02
51– up	13	46.78	6.94	38	50.79	9.38	13	47.04	8.90	38	51.27	10.00
	Fine Arts											
19–22	46	49.66	10.90	82	54.40	10.00						
23–26	195	47.55	8.39	209	54.23	10.20						
27–30	178	47.20	8.74	158	51.62	10.90						
31–34	88	46.04	9.05	110	51.79	10.75						
35–38	66	45.81	8.34	129	50.85	10.10						
39–42	48	46.91	8.24	96	49.27	9.60						
43–46	26	49.00	10.00	60	52.94	9.50						
47–50	12	42.72	8.74	38	45.31	7.49						
51– up	13	49.97	10.30	37	50.56	9.25						

* University of Georgia norms (N = 1716).

TABLE III—ANALYSIS OF VARIANCE FOR AGE DIFFERENCES IN G.R.E.PROFILE SCORES BY TYPE OF UNDERGRADUATE TRAINING

Source of Variation	Mean Square	df	F	p
Chemistry				
Curriculum	13,436.5123	1	154.2557	0.01
Age	463.3162	8	53.1902	0.01
C×A Interaction	319.1813	8	3.6643	0.01
Error	87.1055	1487		
Physics				
Curriculum	3,613.3745	1	39.1493	0.01
Age	822.6669	8	8.9132	0.01
C×A Interaction	266.5429	8	2.8879	0.01
Error	92.2972	1602		
Biological Sciences				
Curriculum	4,577.3940	1	54.7556	0.01
Age	1,091.1333	8	13.1096	0.01
C×A Interaction	1,556.8788	8	18.7054	0.01
Error	83.2315	1601		
Mathematics				
Curriculum	124.5435	1	1.2716	NS
Age	269.6104	8	2.7527	0.01
C×A Interaction	236.2421	8	2.4121	0.05
Error	97.9423	1247		
Social Science				
Curriculum	7,974.3370	1	83.6694	0.01
Age	108.8609	8	1.1422	NS
C×A Interaction	53.3427	8	0.5597	NS
Error	95.3077	1599		
Literature				
Curriculum	15,804.1251	1	791.7060	0.01
Age	270.3324	8	3.0234	0.01
C×A Interaction	53.6651	8	0.6002	NS
Error	89.4138	1599		
Fine Arts				
Curriculum	9,227.8209	1	99.9240	0.01
Age	550.2906	8	5.9589	0.01
C×A Interaction	103.4329	8	1.1200	NS
Error	92.3484	1573		

G.R.E. performance are not significantly related to the type of the students' undergraduate training. (3) Variations in rate of achievement loss with age are not significantly different for students with dissimilar undergraduate training. Table III gives the results of repeating this analysis for each of the seven profile tests.

On the basis of evidence presented in Table III the first hypothesis must be rejected for six of the seven profiles. Only in the case of social science is age not a significant determiner of G.R.E. scores. For the other six subject-matter areas, age increments are accompanied by a significant decrease in mean test scores.

Hypothesis two is untenable in six of seven instances. Curriculum differences are significant for all profiles except mathematics. Students with different types of undergraduate training differ in their retention of subject matter content evaluated by the G.R.E. profiles. Initially B.S. students test significantly higher in chemistry, physics, and biology than non-science majors; however, these differences are attenuated beyond age thirty-nine. Younger students with undergraduate A.B.s were superior on the social science, fine arts, and literature profiles but these differences, too, were less marked in the later age groups.

Age-curriculum interaction terms for the various profiles deserve special comment. Curriculum-age interaction values were not significant for social science, literature, and fine arts. This in part confirms hypothesis number three. For chemistry, physics, and biology, however, interaction is significant; that is, the superiority or inferiority of the two curriculums is more pronounced at certain age levels than at others. The rate of science achievement loss with age is not the same for both science and non-science undergraduate majors.

DISCUSSION AND SUMMARY

In previous studies which reported significant G.R.E. achievement losses with age it was suggested that recency of undergraduate training might be a factor determining these marked changes in achievement with increasing age. There is little evidence in the present study to support this argument. Four years after graduation all significant chemistry and physics achievement differences have been erased; most biological science differences have

also disappeared. On the non-science profiles, recency of undergraduate training is not significant at any age level.

The type of the student's undergraduate training is reflected in his achievement scores at all age levels and in the expected directions. Students who majored in the social sciences and humanities were superior on profiles representing their fields while science majors excelled on the chemistry, physics, and biology tests. Sharp, positively accelerated losses in mean scores for science majors on the science profiles almost wipe out their initial superiority by age thirty-nine.

The full influence of curriculum differences on achievement variations is not clear from the data reported in the previous section. It may be said, however, that the rate of achievement loss with age is differential for students with different types of training but after age thirty-nine most curriculum effects are insignificant.

IMPLICATIONS

Implicit in the experimental design of the Graduate Record Examination were two hypotheses: first, that graduate level subject-matter achievement tests could be used efficiently by deans and admission officers for the selection and evaluation of graduate students; two, that these tests would reveal to the students their strengths and weaknesses in broad areas of learning covered by the profiles. The original G.R.E. standardization procedure further implied that all graduate students regardless of age or type of undergraduate training could be evaluated in terms of the same "standard coin."

Selection and Evaluation. Admission officers using the G.R.E. for selection purposes should be cognizant of the differential decline of profile scores with age and type of previous training. Scores become much more meaningful when comparisons are made between individuals from the same chronological age group. In some areas there are influences that operate to favor the younger student, some areas are apparently unaffected by age changes, while some influences favor the older age group.

Counseling and Guidance. The individual student who uses G.R.E. profiles for evaluating his past learning experiences and for planning his educational future should not only be aware of

his relative standing among all graduate students but also he should know how he compares with age peers who have had the same type of undergraduate training. Recency of undergraduate training can not be entirely ignored in educational planning at the graduate level but that is not as significant a determiner of G.R.E. profile scores as age or type of undergraduate training.

REFERENCES

1) R. Travis Osborne and Wilma B. Sanders, "Comparative Decline of Graduate Record Examination Scores and Intelligence With Age," *The Journal of Educational Psychology*, Vol. 45, No. 6: October, 1954, 353–358.

2) R. Travis Osborne and Wilma B. Sanders, "Differential Decline in Graduate Record Examination Scores With Age," *Journal of Genetic Psychology* (in press).

3) R. Travis Osborne and Wilma B. Sanders, "Variations in Graduate Record Examination Performance By Age and Sex," *Journal of Gerontology*, Vol. 9, No. 2: April, 1954, 179–185.

4) James E. Wert, Charles O. Neidt and J. Stanley Ahmann, *Statistical Methods in Educational and Psychological Research*. New York, Appleton-Century-Crofts, Inc., 1954, 435 pp.

THE DEVELOPMENT OF NEW SCALES FOR THE APTITUDE AND ADVANCED TESTS OF THE GRADUATE RECORD EXAMINATIONS[1]

MARGARET K. SCHULTZ and WILLIAM H. ANGOFF

Educational Testing Service, Princeton, N. J.

The Graduate Record Examinations offer, among other tests, an Aptitude Test, yielding a Verbal score and a Quantitative score, and a series of Advanced Tests in the usual major areas of college study—biology, chemistry, history, psychology and so forth. These tests are taken primarily by college seniors and graduate students in testing programs sponsored by their colleges, and by applicants to graduate schools which require or recommend the tests for admission.

The scaled score system for these tests and for the other tests in the Graduate Record Examinations program is one which defines a scaled score of 500 as the mean score for the particular standardization or reference group on which the scale is based, and 100 as the standard deviation of these scores. Such a score system has little inherent normative value, however, if the standardization group whose performance is the basis for the scale is not reasonably representative of groups who currently use the tests. In the case of the Aptitude Test, groups now being tested are far more heterogeneous and generally lower in ability than the original standardization group, which was composed of first-year graduate men tested at four eastern universities in 1941. This change in the characteristics of the groups being tested was one reason for the development of a new scale for the Aptitude Test. A second reason was the fact that the Aptitude Test itself had undergone revisions in content and scope since the original scaling. It might be noted that the Aptitude Test, as such, was introduced in 1949. Prior to that time, the Verbal and Quantitative scores were obtained as two parts of the Profile Tests.

Aptitude Test scores were first reported on this new scale in the fall of 1952. At the same time, new scales for seventeen of the

[1] Paper presented at the 1953 meeting of the American Psychological Association.

Advanced Tests were also introduced. There were two reasons for re-scaling the Advanced Tests. For one thing, fourteen of them had been extensively revised since the previous scales had been established, and the allotted testing time had been extended from one hour and forty-five minutes to three hours. Secondly, it seemed advisable to make a change in the type of score scale used for the Advanced Tests. The previous scales, established in 1947, had been so set that 500 was the mean scaled score on each particular test for a group of seniors majoring in the field covered by that test, and the standard deviation of their scores was 100. A different standardization group was therefore used for each test, with no adjustment made for group differences in ability. Now it is generally found that some fields of specialization attract more able groups of students, while other fields of specialization attract less able students. Also, some fields attract a more heterogeneous group of students than do others. As a result of this kind of self-selection, the groups of candidates taking different Advanced Tests differ markedly from one another with respect to level and range of ability. In the development of the new Advanced Test score scales, it was decided to design them to reflect these differences in ability as measured by performance on the Aptitude Test. It should be emphasized that both this new type of scale and the 1947 scale preserve the rank order of raw scores on a test and therefore both yield the same percentile ranks for a given group of students.

The data for the re-scaling of the Aptitude and Advanced Tests were collected in the spring of 1952. A single standardization group was chosen. This group consisted of two thousand and ninety-five graduating seniors, six hundred and eighty-six women and one thousand four hundred and nine men, in eleven colleges. The colleges were selected in such a way as to yield a fairly wide range of scores and a level of performance near the average of the groups who normally take the tests. In addition, it was judged that the colleges chosen were widely enough known so that the performance of their students could be evaluated by other test users. It was felt that this knowledge would add to the normative value of the scores themselves. The colleges in the standardization group have been listed in the norms bulletins for the tests.

Each student in the standardization group took both parts of the Aptitude Test—that is, Verbal and Quantitative. The new scale for each part of the Aptitude Test was established by a linear

transformation of the raw scores obtained by the *entire* group of two thousand and ninety-five seniors so as to yield scaled scores with a mean of 500 and a standard deviation of 100. In the scaling of the Advanced Tests, a more complex statistical procedure was employed in order to yield scales which would reflect the performance of the particular Advanced Test group on the Aptitude Test. In addition to taking the Aptitude Test, each of the two thousand and ninety-five seniors also took the Advanced Test in his major field of study. For each Advanced Test subgroup, regression coefficients were determined for predicting Advanced Test scores from Verbal and Quantitative Aptitude Test scores. Estimations were then made of the raw score mean and variance of the *entire standardization group* on each Advanced Test, utilizing the regression coefficients and also the measured differences between the entire group and each Advanced Test subgroup on the two Aptitude Test scores. The equations used in this type of score scaling were developed by Dr. Ledyard R. Tucker of Educational Testing Service (see pages 293–4).

Once estimates of the raw score mean and variance of the entire standardization group on each Advanced Test had been obtained, linear transformations were determined which would result in a scaled score mean of 500 and a standard deviation of 100 on each test for the entire standardization group of two thousand and ninety-five seniors.

Figure 1 shows in graphical form the level and dispersion of scores for each Advanced Test subgroup on the two parts of the Aptitude Test and also on the Advanced Test taken by that group. The total group has, by definition of the new scale, a mean of 500 and a standard deviation of 100 on each part of the Aptitude Test. It can be seen that the Advanced Test subgroups differ rather markedly from one another, and from the total group, in the Aptitude Test scores. For example, as may be seen in Table 1, there is a difference of seven points in the mean Verbal score and sixty-two points in the mean Quantitative score between the Chemistry subgroup and the entire standardization group. It is this kind of difference—the differential performance of each Advanced Test subgroup on the Aptitude Test—which has been taken into account in establishing the Advanced Test scales, and reflects itself in a high mean score (530) for the Chemistry subgroup on the Advanced Test in Chemistry.

FIGURE 1

MEANS AND STANDARD DEVIATIONS OF SCALED SCORES ON THE GRE APTITUDE AND ADVANCED TESTS, OBTAINED BY EACH ADVANCED TEST SUBGROUP IN THE STANDARDIZATION PROGRAM

(Mean score is indicated by a horizontal line drawn through the center of each bar; the standard deviation is indicated by the length of the bar above and below the mean)

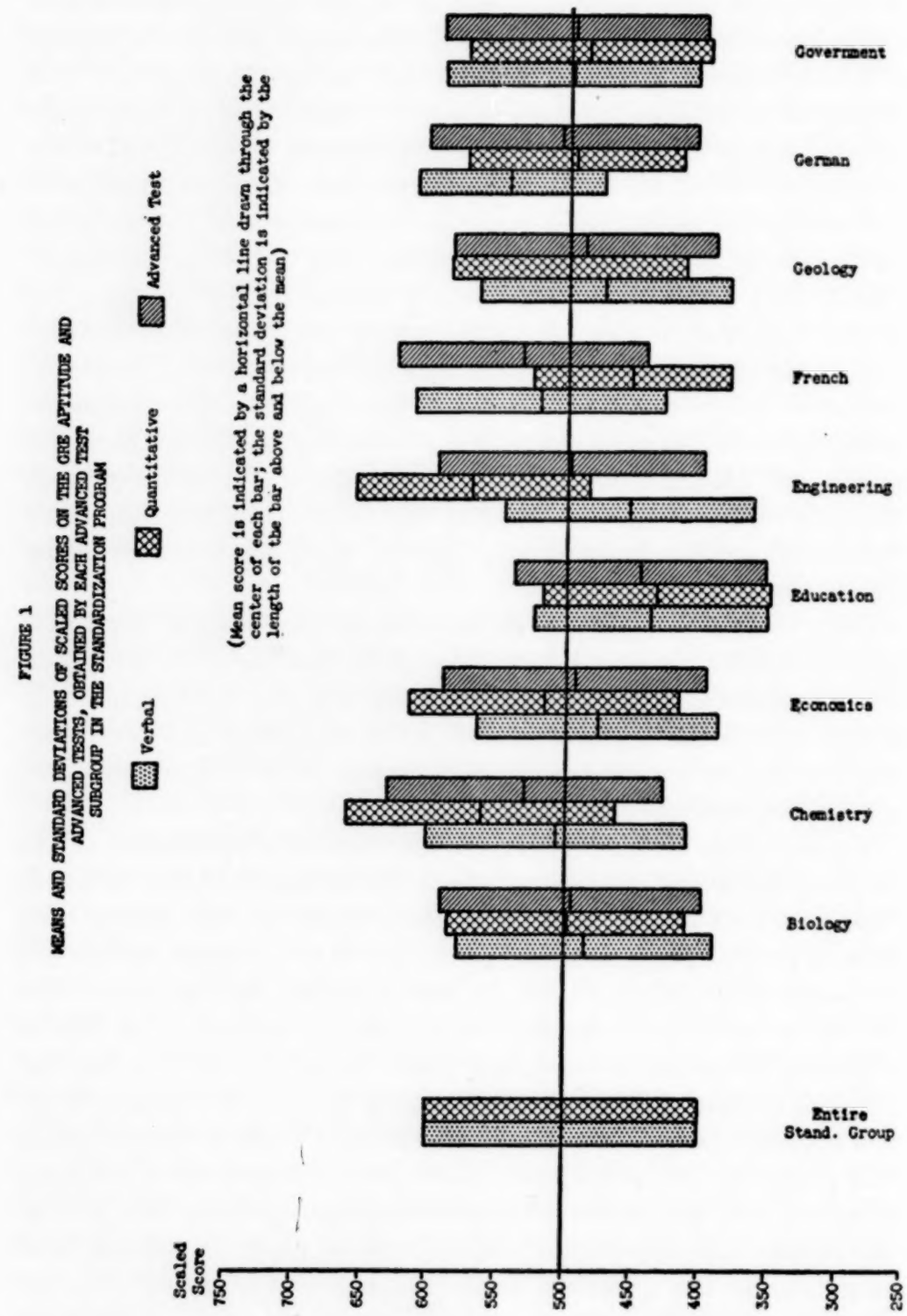

FIGURE 1 (Continued)

MEANS AND STANDARD DEVIATIONS OF SCALED SCORES ON THE GRE APTITUDE AND ADVANCED TESTS, OBTAINED BY EACH ADVANCED TEST SUBGROUP IN THE STANDARDIZATION PROGRAM

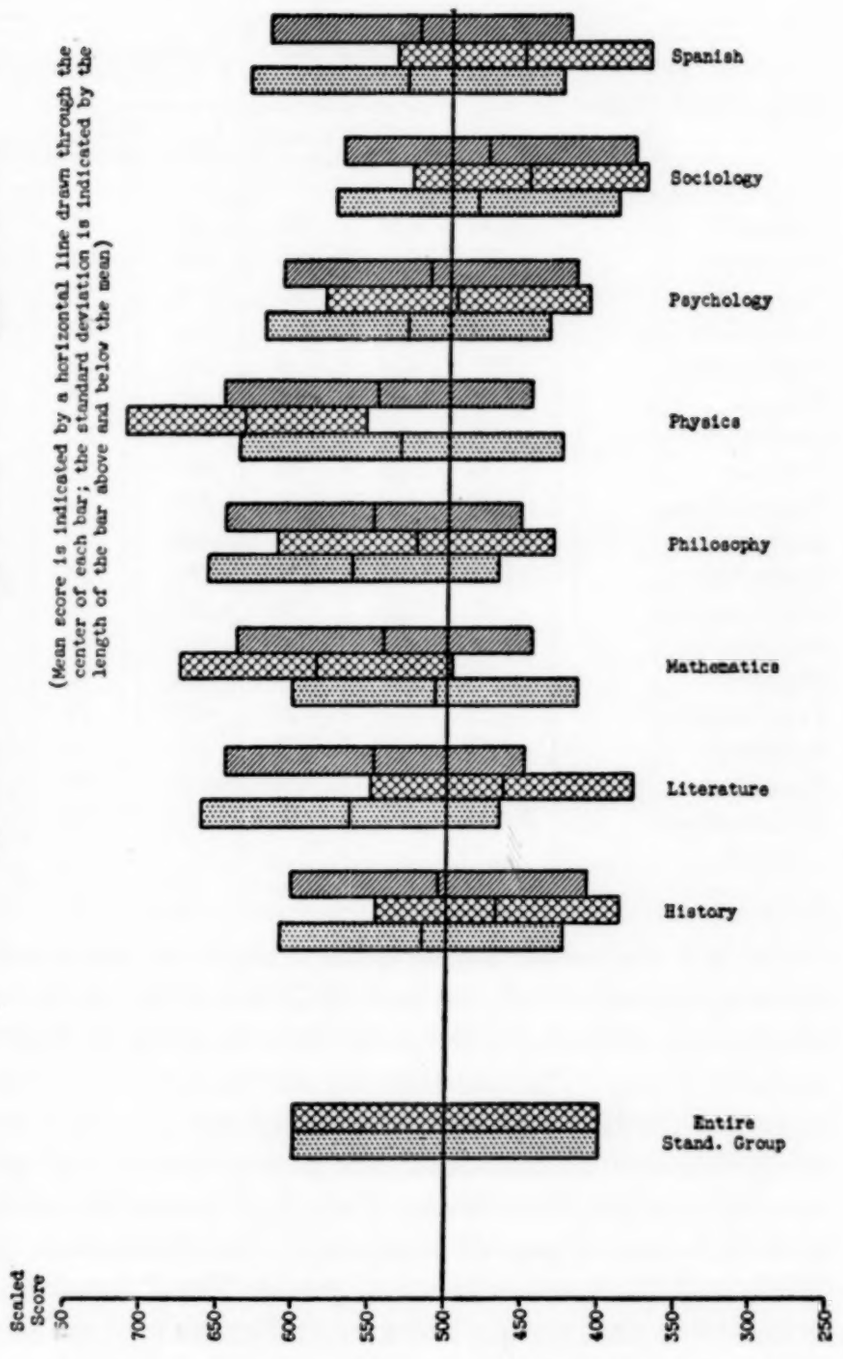

TABLE I—SCALED SCORE MEANS AND STANDARD DEVIATIONS ON APTITUDE AND ADVANCED TESTS, BY ADVANCED TEST SUBGROUP

Advanced Test Subgroup	N	Aptitude Test-Verbal		Aptitude Test-Quantitative		Advanced Test	
		Mean	Standard deviation	Mean	Standard deviation	Mean	Standard deviation
Biology	209	486	94	499	87	495	96
Chemistry	180	507	96	562	99	530	101
Economics	239	476	89	516	99	494	97
Education	180	438	86	434	83	446	93
Engineering	151	454	92	570	86	497	98
French	32	520	92	453	72	533	92
Geology	35	473	92	500	86	488	97
German	10	543	69	495	79	505	99
Government	146	498	93	485	89	496	97
History	181	517	93	468	80	506	97
Literature	239	564	98	463	86	548	99
Mathematics	81	508	93	587	90	542	97
Philosophy	31	563	96	521	90	549	97
Physics	49	531	106	633	79	546	101
Psychology	171	527	94	495	86	512	96
Sociology	127	482	93	447	77	474	96
Spanish	34	529	102	451	83	520	99
Entire stand. group	2,095	500	100	500	100	—	—

Table 1 shows the scaled score means and standard deviations on the Aptitude and Advanced Tests obtained by each of the subgroups—essentially the same data as given in Figure 1, but in numerical form. The fact that the Advanced Test mean and standard deviation for any particular subgroup are above or below 500 and 100 means that the Aptitude Test scores for that group differed correspondingly from the Aptitude Test scores for the entire group, and that some degree of correlation existed between Aptitude and Advanced Test scores for that group. The Advanced Test means range from 446, for the Education Test subgroup, to 549, for the Philosophy Test subgroup. The standard deviations range from 92 to 101. By way of contrast, the 1947 scaling method for the Advanced Tests would have yielded a mean of 500 and a standard deviation of 100 for each of these groups. In the new scales, the 500-point on each Advanced Test is the estimated average score

TABLE II—CORRELATIONS AMONG SCORES FOR EACH ADVANCED TEST SUBGROUP

Advanced Test Subgroup	N	Correlations			
		Aptitude test verbal vs quantitative	Aptitude verbal vs advanced test	Aptitude quant. vs advanced test	Multiple R: adv. test vs apt. verbal, quant.
Biology	209	.47	.50	.58	.63
Chemistry	180	.57	.38	.53	.53
Economics	239	.44	.57	.51	.64
Education	180	.57	.74	.50	.75
Engineering	151	.55	.58	.48	.61
French	32	.56	.54	.10	.60
Geology	35	.50	.57	.52	.63
German	10	−.09	.09	−.27	.28
Government	146	.47	.69	.45	.71
History	181	.49	.61	.38	.62
Literature	239	.49	.82	.45	.82
Mathematics	81	.29	.20	.46	.47
Philosophy	31	.49	.75	.22	.77
Physics	49	.50	.62	.44	.64
Psychology	171	.38	.60	.46	.65
Sociology	127	.53	.67	.53	.70
Spanish	34	.66	.31	.12	.33
Entire stand. group	2,095	.39	—	—	—

on each test of the *entire* standardization group, not the obtained average score of any particular subgroup. In the 1947 scales, the 500-point on each Advanced Test was the *obtained* average score for the particular group of seniors majoring in the field of the test.

As indicated above, the extent to which the level and range of Aptitude Test scores for a particular Advanced Test subgroup are reflected in the Advanced Test scale depends upon the correlation of each part of the Aptitude Test with the particular Advanced Test. Table 2 shows these zero-order correlations and the multiple correlations. For the total group of two thousand and ninety-five seniors, the correlation between the two parts of the Aptitude Test was 0.39. All but four of the multiple correlations are above 0.60, although only four are above 0.70. In the case of two Advanced Tests, particularly German and Spanish, the multiple correlations were extremely low—0.28 and 0.33. For those tests, it would be

expected that the scaling is relatively unreliable. In the extreme case, where the multiple correlation is zero, the best estimate of the performance of the total group on the particular Advanced Test is the performance of the subgroup (taking the Advanced Test) itself. Since it is these estimates which are arbitrarily translated to a mean and standard deviation of 500 and 100 respectively, the type of scale resulting from a multiple correlation of zero would be identical to the one previously used for the Graduate Record Examinations. That is, it would be one where selective factors affecting the level and range of general verbal and quantitative ability (as measured by the Aptitude Test) would essentially be disregarded in setting the separate Advanced Test scales, since knowledge of that performance does not aid in the estimation process.

It should be pointed out that in the case of six of the Advanced Tests the numbers of cases for the subgroups were less than fifty. In the case of German this number was only ten. Obviously, then, these scales are extremely unreliable. However, since the primary effort in setting these scales was to determine the performance of a generally representative total standardization group, the distribution of cases among the various major fields of study was accepted as it was found. To add more cases to individual subgroups and thus to change the proportions in the separate major field groups arbitrarily would have meant changing the nature and composition of the total standardization group, which, in turn, yielded the best estimate of the proportions in the populations. At any rate, it was felt that since the unreliable scales involved tests that were taken by relatively few examinees, extensive harm would not be done.

In summarizing the characteristics of the new GRE scales, two main points can be made. First, the new scale for the Aptitude Test incorporates normative data for a group of students which is clearly identified and fairly representative of a large portion of the test users. Second, the new scales for the Advanced Test are adjusted for the differences among the groups in Aptitude Test performance. The scales for all tests are therefore based upon the performance of a single, over-all standardization, or reference group. This type of scale system reduces a major source of possible error in interpretation that was present in the old-type scales

where each test was separately standardized on a different group of students.

Under the old scale system developed in 1947, an individual's Advanced Test score indicated his standing relative to the group of people on whom that particular test was standardized. Since the separate test-groups were different (because of self-selective factors) the scale referents for the various Advanced Tests necessarily differed one from another. Under the present scale system, an individual's Advanced Test score indicates his standing relative to the estimated performance of the entire standardization group. Since this group is constant for all tests, it therefore constitutes a single common referent for all the Advanced Test scales.

SUMMARY

The new scales for the Aptitude and Advanced Tests of the Graduate Record Examinations have been developed with the use of data collected on two thousand and ninety-five second semester seniors in eleven colleges which are generally representative of the schools that normally administer the tests to their students. Scores on the Verbal and Quantitative parts of the Aptitude Test were converted to yield scaled score means of 500 and standard deviations of 100. Scores on the Advanced Tests were adjusted to reflect differences in Aptitude Test scores among the groups taking the separate Advanced Tests. It is considered that this type of score scale would lead to fewer errors of interpretation than the previous scales, in which raw scores on each of the Advanced Tests were converted directly to a scale with a mean of 500 and standard deviation of 100.

Equations Used for the Scaling of the GRE Advanced Tests

Notation:

v, q = scaled scores on the Verbal and Quantitative parts of the Aptitude Test
x = raw scores on the Advanced Test
t = entire standardization group
s = subgroup taking the particular Advanced Test
$\tilde{M}, \tilde{\sigma}^2$ = estimated mean and variance
M, σ^2 = obtained mean and variance

b = regression coefficient
C_{vq} = covariance between Verbal and Quantitative Aptitude Test scores

$$\tilde{M}_{x_t} = M_{x_s} + b_{xv \cdot q_s}(M_{v_t} - M_{v_s}) + b_{xq \cdot v_s}(M_{q_t} - M_{q_s})$$

$$\tilde{\sigma}^2_{x_t} = \sigma^2_{x_s} + b^2_{xv \cdot q_s}(\sigma^2_{v_t} - \sigma^2_{v_s}) + b^2_{xq \cdot v_s}(\sigma^2_{q_t} - \sigma^2_{q_s}) + 2b_{xv \cdot q_s} b_{xq \cdot v_s}(C_{vq_t} - C_{vq_s})$$

Linear Transformation to Obtain Scaled Scores (y):

$$y = Ax + B,$$

where

$$A = \frac{100}{\tilde{\sigma}_{x_t}} \quad \text{and} \quad B = 500 - A\tilde{M}_{x_t}$$

THE LEARNING OF SUBJECT MATTER IN TEACHER-CENTERED AND GROUP-CENTERED CLASSES

GERARD V. HAIGH[1] and WARREN SCHMIDT

The Menninger Clinic and the University of California

A perennial problem for the teacher is that of evaluating teaching methods. A contemporary trend in education involves experimenting with shifting control in the classroom from the teacher to the students as a group. The many variations in the control pattern which have been tried may be loosely classified as teacher-centered and as group-centered. One critical question asked is: "Which approach results in greater student mastery of the subject matter?" Existing studies which bear on this question have yielded conflicting answers. Studies by Asch (*1*), Husband (*8*) and Guetzkow, *et al.* (*6*) show small but significant differences favoring the teacher-centered approach. A study by Faw (*5*) shows a difference favoring the group-centered approach. Studies by Landsman (*9*), Smith and Johnson (*10*), Eglash (*4*) and Wispé (*11*) show no significant differences between these two approaches. This conflicting evidence suggests the need for further isolation of the pertinent variables.

One factor confounding the results of the above studies is the almighty importance of "The Grade" to goal-straining American students. If getting a good grade depends on acquisition of knowledge, then this grade motivation may induce the acquisition of knowledge independent of teaching method. In all of the above studies indicating no significant difference between teaching methods, it was necessary for the students to do well in examinations in order to get a good grade. In fact, in all but two, the mark on the final examination or the final grade itself was used as the criterion for the amount of knowledge acquired. The two exceptions (*9*, *10*) used achievement tests but also had course examinations. We can probably infer positive transfer from studying for the regular examinations to performance on the achievement tests.

The one study suggesting superior results for the group-centered

[1] This study was done while the authors were teaching at Springfield College, Springfield 9, Massachusetts.

type of class (*5*) also used the course examination as a criterion. Since student grades were apparently dependent upon this criterion, the results involve the same ambiguity as those studies reporting no difference.

Only in the study reported by Asch (*1*) were student grades in the group-centered section *not* determined by examinations. Students were required to determine their own grades and were told that the final examination did not count toward their grade. In this study, the teacher-centered section did significantly better on a final objective examination.

These findings suggest that if students in a group-centered class are not "required to learn" via the grade-examination system, then they do not learn as much information as do students in a teacher-centered class.

However, even if we accept this finding from Asch's study, it does not rule out the possibility that some students may learn more in a teacher-centered class while others may learn more in a group centered class. The study by Wispé (*11*) highlights the importance of interaction between student personality and teaching method. In all the foregoing studies, students were arbitrarily assigned to experimental and control groups and the results treated as group data. How would consideration of personality differences affect their findings?

The present study is one more attempt to answer the question as to whether teacher-centered classes or group-centered classes are preferable for the learning of subject matter. It is similar to Asch's study in avoiding the influence of grade-striving by incorporating a self-grading system in the group-centered sections. It pushes beyond the Asch study into the area of student personality differences by assigning students to the two types of sections on the basis of their stated preference. Basically, this study addresses itself to the following question: When students are permitted to choose either a teacher-centered class or a group-centered class and when the students in the group-centered class are not required to submit knowledge to the test of the examination-grade system, will there be a difference between these two types of classes as to the amount of subject matter learned?

At the time of this study (1952), Springfield College had a three-course sequence in psychology, consisting of general, child, and

adolescent psychology, required of all sophomore students. On the first day of the second course, students were presented with the following outline describing two teaching methods:

Assumptions Involved in Two Plans for Teaching Psychology 2

Plan A (Teacher-centered)	Plan B (Group-centered)
Curriculum	
The instructor knows the facts which are important for the student to learn.	The students have questions about the subject.
It is the instructor's job to get the students interested in learning these facts.	It is the job of the class (including the instructor) to work together in finding answers to these questions.
Classroom Method	
Communication: Most of the time, the instructor talks to the students. Sometimes the students talk to the instructor.	Everybody talks with everybody.
Students listen mostly to the instructor.	Students listen to each other as well as to the instructor.
Rôle of instructor: Initiates the questions, assigns readings, presents the course material, evaluates students on their learning.	Moderates group discussion, contributes information as needed suggests readings when requested, participates as an active group member.
Rôle of student: Learns the course material as presented by instructor. Does the readings as assigned. Gives evidence of understanding on examinations.	Initiates the questions, contributes his opinion, seeks additional evidence, contributes evidence to group, helps arrive at conclusions, helps others contribute to group.
Evaluation: Instructor assumes full responsibility for gathering evidence and assigning grade.	Students assume responsibility for providing evidence. Instructor and student share responsibility for determining grade.

After general discussion of these two methods, each student filled out a brief questionnaire designed to help him weigh the relative merits of each method. At the end of the questionnaire the student stated which method he preferred for the course. The class was organized so that each student's preference was met. One instructor

taught three teacher-centered sections. Another instructor and two graduate assistants[2] each led a group-centered section.

Twelve weeks later, on the first day of the third course, students were again asked to choose their preferred method, with freedom to change if they desired. At the end of this third course, all sections were given the Horrocks-Troyer Test (*7*). The examination was cut in half by using every other question in order to ameliorate the effect of fatigue on students in the teacher-centered sections who had just finished a final examination in the same sitting. All students were informed that this test would not affect their course grade.

POPULATION

The subjects in this study consisted of one hundred and ten students who chose one method and followed it through the two courses. An additional fifty students were involved in the original choosing but do not appear in the study either because they changed methods or because they did not immediately follow up the second course with the third course. As reported previously by Ashmus and Haigh (*3*), in the original choosing, eighty-one students preferred a teacher-centered class while seventy-nine preferred a group-centered class. Of the students continuing in a class of their original choice, there were fifty-five in each group. As reported previously (*3*), the groups originally choosing each method displayed no significant differences with respect to age, sex, academic aptitude, previous academic achievement or major in college. It is assumed that this equating persists for the whittled-down groups.

While comparability seems to exist within the two groups with respect to general academic achievement, it might be possible that they differed significantly with respect to knowledge of psychology at the outset of the study. To test for this possible biasing factor, the groups were compared with respect to the marks which the members of each obtained on the mid-term examination in Psychology 1.[3] A resultant T-score of 0.59 shows no significant dif-

---

[2] We wish to thank Mable Ashmus and Richard Wiseman for serving as group leaders.

[3] The final grade was not used as a criterion because it was determined largely on the basis of term projects which the students were permitted to submit in any quantity they preferred. Thus, the grades came to be influ-

TABLE I—MEAN SCORES ON OBJECTIVE EXAMINATION OF TEACHER-CENTERED CLASS AND GROUP-CENTERED CLASS

	Teacher-Centered Class	Group-Centered Class
Mean score on objective exam.	25.6	24.7
S.D. of distribution	5.89	6.2

C.R. of the difference t = 0.78.

With one hundred and eight degrees of freedom, the level of significance is between 0.4 and 0.5.

ference between the experimental and the control group re knowledge of psychology prior to the experiment.

RESULTS

Table I gives the results of this study.

The difference between these two groups in the learning of subject matter is clearly not significant.

The pattern of results among the six sections was the same as that for the experimental and the control groups. There were no significant differences between any two sections.

DISCUSSION

Of the studies cited in the introduction, this one is directly comparable only with that of Asch in that both involve grade determination in the group-centered sections independent of any examinations and so avoid this influence upon knowledge-gain. In the Asch study, the teacher-centered section did significantly better on the two parts of a final objective test, suggesting more knowledge-gain with this type of teaching method.[4]

The present study suggests that this advantage of the teacher-centered method is nullified when students are permitted to choose the type of method which they prefer. Probably this choice is a reflection of personality factors which also influence the capacity of students to respond to these two types of teaching methods.

---

enced as much by diligence as by knowledge. Incidentally, the mass of paper accruing discouraged the instructors from ever using this system again.

[4] Asch also reports "supplementary evaluation data indicate that the nondirective (group-centered) students gained a greater knowledge of diversified subject matter and did more independent reading and thinking about psychology than did the control (teacher-centered) subjects."

(Choice might merely be a reflection of previous experience but this seems to have been ruled out in the present study. Although among those students who had had no experience previously with the group-centered method there was a significant preference for the teacher-centered method, among those who had had experience with both there was no significant preference for either). (*3*) Wispé's study (*11*) indicates that personality variables are associated with differences in attitude toward directive and permissive teaching. He found that students whom he described via projective tests as "personally insecure" had particularly unfavorable attitudes toward permissive teaching; those whom he described as "independent" preferred permissive teaching; and those whom he described as "satisfied" reacted favorably to both methods. In a corollary study to the present one, Ashmus, Bedrosian and Wiseman (*2*) found the following personality differences between students choosing these two methods:

(*1*) Students choosing a group-centered class are more flexible in problem-solving as measured by the Einstellung Test (Wiseman).

(*2*) Students choosing a group-centered class were better able to cope with inconsistencies and ambiguities in reproducing stories (Bedrosian).

(*3*) Students choosing a teacher-centered class were more stereotyped and had less insight in describing their self-picture (Ashmus).

The differences found in the above studies were small; some of them were not statistically significant, but all were consistent. They suggest on the one hand the picture of a student who is flexible in his thinking, able to cope with inconsistencies and ambiguities, understanding and accepting of himself and probably able to enjoy and profit from the give and take of group discussion. On the other hand they suggest the picture of a student who seeks the definite, the concrete, and the ordered, avoiding ambiguity and inconsistency, seeing himself (and life) in stereotyped ways and who is probably able to gain most in satisfaction and knowledge from a teacher-centered class. It is probably in some such vein as this that the results of the present study may best be interpreted.

CONCLUSION

Evidence is presented to show that when students are permitted to choose between a teacher-centered class and a group-centered

class and when the students in the group-centered class are not required to learn subject matter by the examination-grade system, there is no significant difference between these two types of classes in knowledge of subject matter at the end of the term.

REFERENCES

1) M. J. Asch, "Nondirective Teaching in Psychology," *Psychological Monographs*, 65: 24, 1951.

2) M. Ashmus, A. Bedrosian and R. Wiseman, *Clearness of Perception as Related to Choice of a Student-Structured Class*. Master's thesis, Springfield College, 1952.

3) M. Ashmus and G. Haigh, "Some Factors which may be Associated with Student Choice between Directive and Nondirective Classes," *American Psychologist*, 7: 247, 1952 (abstract).

4) A. Eglash, "A Group Discussion Method of Teaching Psychology," *Journal of Educational Psychology*, 45: 257–267, 1954.

5) V. Faw, "A Psychotherapeutic Method of Teaching Psychology," *American Psychologist*, 4: 104–105, 1949.

6) H. Guetzkow, E. L. Kelly and W. J. McKeachie, "An experimental Comparison of Recitation, Discussion, and Tutorial Methods in College Teaching," *Journal of Educational Psychology*, 45: 193–207, 1954.

7) J. Horrocks and M. Troyer, *A Test for Knowledge of Facts and Principles in Adolescent Development*. Syracuse University Press, 1946.

8) R. W. Husband, "A Statistical Comparison of the Efficacy of Large Lecture Versus Smaller Recitation Sections upon Achievement in General Psychology," *Journal Psychology*, 31: 297–300, 1951.

9) T. Landsman, *An Experimental Study of a Student-Centered Learning Method*. Unpublished Ph.D. thesis, Syracuse University, 1950.

10) H. C. Smith and D. Johnson, "An Experimental Study of Attitudes and Achievement in the Democratic Classroom," 1951 meetings of A.P.A.

11) L. G. Wispé, "Evaluating Section Teaching Methods in the Introductory Course," *Journal of Educational Research*, 45: 161–186, 1951.

A STUDY OF THE FAMILY BACKGROUND OF THE GIFTED

WALTER B. BARBE

University of Chattanooga, Chattanooga, Tennessee

Of what influence has heredity and environment been in the development of the gifted child? Specifically, how does the family background of the gifted child differ from that of the child who is average in intelligence? Unfortunately, it is not possible to determine from such a study as this whether heredity or environment played a larger part in the development of the superior intellect of the subject. But by means of such a study, factors in the development of the gifted may be revealed. In a follow-up study of a group of gifted Negro subjects, Jenkins concluded:

> . . . desirable as it would be to know which of these factors (heredity or environment) has been most potent in the development of our subjects, the writer is unable to present any crucial data on the question. The data relative to heredity are meager, and to some extent, superficial; and while the picture of the cultural background is more complete, even here there are intangible factors which elude objective evaluation (*1*).

The 456 subjects in this study received an I.Q. of 120 or above on the 1917 Form of the Stanford-Binet. These data were taken from the records of the Psychological Clinic of the Cleveland Board of Education. The range of I.Q. is from 120 to 164, with a mean I.Q. of 130.2. The largest number of subjects were in the 125–129 range (37.3 per cent), while almost sixty-two per cent were between 125 and 135. This placed all of the subjects in about the upper ten per cent (*2*) of the population of the United States at the time they were tested. A large percentage of the group (forty-four per cent) were in the upper one per cent of the population in intellectual ability as measured by this particular test.

All of the subjects were graduates of the Major Work Program of special classes for public school children in Cleveland, Ohio. The data for this study were obtained from information reported on a five-page printed questionnaire which was distributed to the graduates of the program over the last fifteen years. Of those who received the questionnaire, a return of seventy-seven per cent was received.

RACIAL AND RELIGIOUS BACKGROUND

It is difficult to determine the racial background of a group of subjects. The question arises just how far back the subject should trace his ancestry. With regard to the ancestry of gifted subjects, Hollingworth states:

> So few data have been gathered to show the proportion of gifted children in relation to race, that it is perhaps scarcely worth while to discuss the topic except to say that we are ignorant of the facts. We have, however, a few studies of the proportion of gifted in samplings of the various races found at present in the United States (*3*, pp. 68–69).

In a study by Witty (*4*), the "racial stock included a preponderance of English, Scotch, German, and Jewish ancestors." Ninety-six per cent of the parents were American-born. These findings are similar to those of Terman and Oden:

> The reports on racial origin indicate that, in comparison with the general populations of the cities concerned, there is about a one hundred per cent excess of native-born parentage, a probable excess of Scottish parentage and a deficiency of Italian, Portuguese, Mexican, and Negro ancestry (*5*).

The racial stock of the subjects in the present study is predominantly German, nearly half (47.6 per cent) of the subjects reported having some degree of German ancestry. The next highest group mentioned, twenty-three per cent, was English. The next most frequent were: Hungarian, 14.7 per cent; Russian, 14.4 per cent; and Polish, 10.2 per cent. Hardly any European country was not at least mentioned.

The fact that the population of Cleveland consists of such a diverse foreign element would tend to make the racial background of the subjects different from the subjects in both Terman and Witty's studies. According to the 1940 Census (*6*) of the City of Cleveland, the largest foreign-born element, about thirteen per cent, was Polish. Czech, Hungarian and Italian each made up about twelve per cent of the foreign-born population. German, Yugoslavian, English and Russian each represented about seven per cent.

Slightly more than two and one-half per cent of the total sample are Negroes. The percentage of Negroes has risen from less than one per cent in 1938 to nearly five per cent in 1952. According to the census, the per cent of Negroes in Cleveland has risen from eight per cent in 1930 to 9.5 per cent in 1940 (*6*).

About 11.3 per cent of the group was Catholic, 46.3 Protestant, and 38.8 per cent Jewish. Other groups make up the remaining 3.8 per cent. As Hollingworth reports of New York City (*3*), children of Catholic parents, many of whom are gifted, are commonly being educated in parochial schools. This explains the rather low percentage of Catholic subjects in Major Work classes. The Jewish group appears to be represented in far greater numbers than its proportional share. This is also true of the gifted in the public schools of New York where Hollingworth reports there is "a marked excess of Jewish parentage (*3*, p. 70)."

ECONOMIC BACKGROUND

An important phase of a study of the gifted which has not received adequate attention is their socio-economic background. This is difficult to determine and, when done in retrospect, is subject to many errors. The procedure followed in this study was to locate the economic tenth of the census tract in which the subject had lived while he was in public school (*7*). This gave an indication of the rent and property value of the neighborhood in which the gifted subject had been reared. The results of this phase of the study are presented in Table I.

The economic tenth from which the greatest number of subjects came was the seventh, while the sixth and seventh economic tenths included more than fifty-eight per cent of the subjects. This indicates that the background of the majority of the subjects in the study may accurately be described as "upper middle-class."

TABLE I—THE ECONOMIC STATUS OF 456 GIFTED SUBJECTS

Economic Tenths	Per Cent
Highest	1.1
Ninth	7.9
Eighth	10.3
Seventh	37.1
Sixth	21.3
Fifth	11.0
Fourth	5.0
Third	3.5
Second	2.2
Lowest	0.7

ORDER OF BIRTH AND SIZE OF FAMILY

It has long been a popular belief that the gifted child is an only child, or, perhaps, has one sibling. In a study of 253 subjects (*8*), Goddard reported half as being first-born and three-fourths as being either first- or second-born. Of the first-born, forty-five (about eighteen per cent) were only children.

Hollingworth reported that the gifted child had few siblings. In a study by Cobb and Hollingworth, fifty-seven gifted children averaged less than one sibling each (*3*). With respect to the order of birth, they found that more than one-half of their subjects were first-born.

In 1940, Terman reported that "the parents of the gifted subjects had produced. . .an average of 3.09 (children) per family. He states that this rate would more than maintain the stock, but "it appears likely that the subjects themselves will not equal the fertility rate of their parents" (*5*, p. 18).

In the present study two questions were asked to determine the size of the family and the order of birth of the gifted child. About 21.8 per cent of the subjects had no siblings, while 42.6 per cent had only one. Almost twenty per cent had two siblings, and seven per cent three.

About twenty-two per cent of the subjects in this study are only children. This is not as large a number as that found by Cobb and Hollingworth, although their finding that "more than half were first-born" (*3*, p. 180) is substantiated by the fact that 52.5 per cent of the subjects in this study were first-born. About twenty-nine per cent are second-born, and only 9.3 third-born. The data indicate that in this group the gifted child was the first-born in a family of two children.

PARENT OR GUARDIAN OF GIFTED SUBJECTS

Eighty-seven and a half per cent of the gifted subjects were reared by their own parents. The next largest group, 7.2 per cent, were reared by only their mothers. Two per cent were reared by their own mother and a step-father, and 1.5 per cent by their own father and a stepmother. The father or foster parents each reared 0.9 per cent of the total number of subjects.

Witty (*9*), in studying one hundred gifted children in Kansas, found that most of their parents were American-born. This was

TABLE II—PERCENTAGE OF PARENTS OF GIFTED SUBJECTS WHO WERE BORN OUTSIDE THE UNITED STATES

Both parents U.S.-born	51.3
One parent foreign-born	21.1
Both parents foreign-born	27.5

not true of the subjects in this study. Slightly less than fifty per cent of the subjects had one or both parents who were foreign-born. This is partially due to the large foreign element in the population of Cleveland.[1] It emphasizes the contribution of the immigrant to the mentally superior groups of the country. Table II presents these data.

Even though more rigid government controls have been placed on immigration, no trend is noted which would indicate that fewer of the subjects have parents who are foreign-born.

EDUCATION OF PARENTS

Hollingworth states (*3*) that the educational level of the parents of gifted subjects is far above the average for their generation. "In the majority of the cases where the gifted child has been born since 1915, both parents are graduates of high school, and in far more cases than in the population at large both parents are college graduates (*3*, p. 180)." Since all of the subjects in the present study were born after 1915, it is interesting to compare Hollingworth's statements with the data obtained for this group.

Of the fathers of the subjects, 38.4 per cent had a grammar school education or less; 33.6 per cent had a high school education; 8.8 per cent trade or business school; and 19.2 per cent had some college. Of the mothers of the subjects, 32.5 per cent had a grammar school education or less; 42.2 per cent had a high school education; 12.3 per cent trade or business school; and 13.0 per cent had some college. The mothers of the gifted subjects on the average appeared to be slightly better educated than the fathers through the high school and business school levels. However, there were more fathers than mothers who attended college.

MARITAL STATUS OF PARENTS

Terman reported (*5*) that until 1922, 5.2 per cent of the parents of his gifted group had been divorced and 1.9 per cent were sepa-

[1] About 20 per cent of the population of Cleveland were foreign-born.

rated. By 1940, the percentage of divorced and separated parents had risen to 13.9.

The data in the present study are not exactly comparable to the results of Terman's study. The information obtained in this study deals with the marital status of the parents while the subject was in public school. It is perhaps comparable to Terman's 1922 data but is definitely not comparable to his 1940 data.

Eighty-eight per cent of the parents of the subjects were living together while the subject was in public school. About 6.3 per cent were either divorced or separated. This is only slightly higher than the report for Terman's 1922 group and is certainly lower than that of the general population. The remaining five per cent consists of cases where one or both of the parents were deceased.

OCCUPATIONAL LEVEL OF PARENTS OF GIFTED SUBJECTS

The occupations of the parents were listed according to the U. S. census classification. The *Dictionary of Occupational Titles* (*10*) was used to classify the occupations into seven distinct groups: professional and managerial; clerical and sales; service; agriculture, fishery, forestry, etc.; skilled; semi-skilled; and unskilled. The subjects were asked the title and description of the father's occupation. It was possible to classify all but a few of the occupations listed. Where descriptions were not given and two classifications were possible, the data were omitted. Classifications of the parent's occupation were made for four hundred and thirty-seven subjects. Three of the remaining nineteen were on government pensions, while the rest gave no response to the question at all. The data are presented in Table III.

Hollingworth reports (*3*) that more than fifty per cent of the

TABLE III—OCCUPATIONAL LEVEL OF PARENTS OF GIFTED SUBJECTS

	Per Cent
Professional and managerial	40.3
Clerical and sales	22.4
Service	3.7
Agriculture, fishery, forestry, etc.	0.2
Skilled	21.5
Semi-skilled	8.2
Unskilled	3.7

children testing above 140 I.Q. have fathers who are professional men or proprietors. The I.Q.'s of the subjects in the present study are not this high, which may partly explain why only about forty per cent of the parents fall into the professional and managerial group. Hollingworth also states (*3*) that half of the remaining fathers are in semi-professional and clerical occupations. This corresponds to the clerical and sales group of the U. S. Census Bureau, and the data for this group agree with the data in Hollingworth's study.

The fact that over thirty per cent of the parents are in the laboring class, and about forty per cent of these are semi-skilled or unskilled, is noteworthy. It indicates that while the majority of gifted children do come from parents of higher occupational status, the laboring class also contributes a sizeable number.

SUMMARY

In this study data were presented concerning the composition of the group being studied and their family background.

(1) The range in I.Q. of the four hundred and fifty-six subjects was from 120 to 164 with a mean I.Q. of 130.2. Almost sixty-two per cent of the group were within the 125–135 range.

(2) Slightly less than fifty-two per cent of the subjects are females; slightly more than forty-eight per cent are males.

(3) Of the total samples, only 2.6 per cent are Negroes.

(4) About thirty-nine per cent of the subjects are Jewish. The Jewish group is represented in far greater numbers than the size of this group in the total population would lead one to expect.

(5) The economic tenth of the census tract in which the subjects lived while in public schools was most frequently the sixth and seventh. This would characterize the gifted child as being upper middle-class.

(6) The gifted child appears to be either an only child or first-born in a family of two.

(7) Eighty-seven and one-half per cent of the gifted subjects were reared by their own parents.

(8) Almost fifty per cent of the subjects had one or both parents who were foreign-born. These data indicate that the group studied is quite unlike other studies of gifted groups. Previously, the gifted child was found to have American parentage. The high percentage

of foreign-born in Cleveland (approximately twenty per cent) partially explains these data.

(9) The education of the mothers of gifted subjects is slightly higher than that of the fathers, even though more of the fathers went to college.

(10) Forty per cent of the parents were in the professional and managerial group, 22.5 per cent in the clerical and sales, and thirty per cent in the laboring class.

The subjects in this study come from about average backgrounds with respect to occupational level, educational level and marital adjustment of their parents. Economically the majority of them come from an upper middle-class group.

REFERENCES

1) Martin David Jenkins, "A Socio-Psychological Study of Negro Children of Superior Intelligence." Unpublished doctoral dissertation, Graduate School, Northwestern University, Evanston, Illinois, p. 53, June, 1935.

2) Merle R. Sumption, *Three Hundred Gifted Children*, p. 6. Yonkers-on-Hudson, New York; World Book Co., 1941.

3) Leta A. Hollingworth, *Gifted Children*. New York, The Macmillan Co., 1926.

4) Paul A. Witty, "A Study of One Hundred Gifted Children," *Bulletin of Education, University of Kansas*, 2: 8, February, 1930.

5) Lewis M. Terman and Melita H. Oden, *The Gifted Child Grows Up*. Stanford, California, Stanford University Press, 1947.

6) U. S. Department of Commerce, Bureau of the Census, Sixteenth Census of the United States 1940, *Population*, 11: 712, Part 5, Washington, D. C., U. S. Government Printing Office, 1943.

7) Howard Whipple Green, *Census Tract Street Index For Cuyahoga County*, fifth edition. Cleveland Health Council, 1951.

8) Henry H. Goddard, *School Training of Gifted Children*, p. 129. Yonkers-on-Hudson, New York, World Book Co., 1928.

9) Paul A. Witty, "A Genetic Study of Fifty Gifted Children," In *Intelligence: Its Nature and Nurture*, Thirty-Ninth Yearbook, Part II, National Society for the Study of Education. Chicago, University of Chicago Press, 1940.

10) Job Analysis and Information Section, Division of Standards and Research, United States Department of Labor, *Dictionary of Occupational Titles, Part I*, second edition. Washington, D. C., United States Government Printing Office, 1949.

DETERMINING THE RELATIONSHIP BETWEEN INDIVIDUAL INTEREST PROFILES AND OCCUPATIONAL FORMS

ALBERT J. KINGSTON, CLAY E. GEORGE and W. PRICE EWENS[1]

The Agricultural and Mechanical College of Texas

One of the real problems which faces vocational counselors who employ the Kuder Preference Record in counseling high school or college students is determining the occupational significance of the obtained interest profile. Although certain counselors are content to utilize the *a priori* suggestions presented in the manual, others have turned to the data on various occupational groups for assistance (*5*). In each succeeding revision of the inventory's manual, Kuder has presented additional profiles or norms which are the result of various studies. It should be noted, perhaps, that these occupational profiles or norms are based upon studies conducted by different users of the inventory and are not the result of unified, systematic research.

This study represents an attempt to investigate the use of these "occupational profiles" when interpreting the interest profiles of a group of college students.

PROCEDURE

The Kuder Preference Record, Form CH, was administered to a group of white, male students enrolled in the Agricultural and Mechanical College of Texas. Two days later the group was given the following instructions: "List in order of preference the five occupations you would most like to enter after college."

The writers hypothesized that students whose listed vocational choices appeared to be consistent, i.e., tended to reflect a similar pattern of interest, would tend to achieve Kuder interest scores which closely resembled the occupational profile or norm for their chosen profession. Similarly it was reasoned that students whose

[1] Thanks are due to A. E. Denton, R. E. Miller and F. E. McFarland, also of the Agricultural and Mechanical College of Texas, for their help in gathering the data for this study.

stated occupational choices reflected a high degree of inconsistency or discrepancy would achieve Kuder scores which less closely resembled their chosen vocation. In the light of the findings of Strong (*6*) and the theories of Carter (*2*) and Bordin (*1*), the hypothesis seemed worthy of study despite the knowledge that individual deviations and lack of refined measurement might preclude verification.

In order to determine whether the stated occupational choices were consistent or not, three counselors were independently asked to sort the listings. A high degree of consistency was found and only those cases on which there was unanimous agreement were used in further study.[2] Eighty cases deemed homogeneous or consistent and forty-five cases felt to demonstrate a lack of consistency were obtained. The following typical cases may illustrate the difference between the two groups. G. E. C., a freshman, aged seventeen, enrolled in Basic Engineering, listed his choices as follows: (1) Architectural Engineer, (2) Electrical Engineer, (3) Petroleum Engineer, (4) Civil Engineer. It is readily apparent that the occupational choices of the student cluster around Computational, Scientific and Mechanical interests at a professional level. In contrast, H. R. H., a nineteen-year-old sophomore majoring in Civil Engineering listed his choices as (1) Civil Engineering, (2) Farmer, (3) Carpenter, (4) Nuclear Physicist. H. R. H.'s choices reflect a wide difference in level of occupation as well as a discrepency in interest areas.

After the stated choices had been sorted into the two categories—consistent or nonconsistent—each category was re-sorted according to the students' major courses of study. As a result, four groups with sufficient numbers for adequate statistical analysis were obtained. These groups included Business Administration, Education, Physical Education and Engineering. Du Mas' coefficient of profile similarity (*3*) was calculated between Kuder's occupational norms and the Kuder profile of each student (except for the Physical Education group for which no applicable norm was available). By use of the Du Mas technique the degree of similarity between the interest profile for each student and the occupational norm would be ascertained. Close similarity between the profiles

[2] It should be noted that many students were unable to list five choices of occupation so only the first four choices were utilized in the sorting process.

of those students whose stated occupational choices displayed consistency and the Kuder norms would lend support to the hypothesis.

One difficulty lay in the selection of the appropriate occupational norm. Such factors as suitability and sufficient number had to be considered. For example: the scale "All Professional Engineers" was employed because the writers' experience had shown that many freshmen and sophomores change their specified engineering major after they complete two years of basic engineering. The use of this category also allowed the study of a greater number of students than did the more refined occupational profiles. Similar considerations dictated the use of the other occupational profiles, "Office Managers and Chief Clerks" and "All Secondary School Teachers."

RESULTS

The calculation of coefficients of profile similarity (r_{ps}) between individual student profiles and occupational profiles failed to reveal differences between students whose vocational choices reflected consistent patterns of interest and those which did not. Not only was no difference noted between either of the student groups and the occupational profile, but when both student groups were combined for further comparison, the profile coefficients yielded a median r_{ps} of only 0.08.

Although the analysis of profile similarity failed to reveal significant differences, some distinctions in the interest scores of the student groups could be discerned. One of the most readily apparent differences was the higher Mechanical and Scientific scores achieved by engineering students. This observation led to further speculation about the effectiveness of using profile similarity techniques with the Kuder Preference Record. The Du Mas technique provides a method of determining profile similarity by analysis of the slope of the various scales which comprise the profile. In this study it appeared that at least two of the Kuder scales—Musical and Artistic—were nonfunctional in the measurement of the interests of any of these specific vocational groups. These scales, however, affected the coefficients of profile similarity as much as some of the more pertinent scales.

In order to check this observation, an analysis of variance was made of the scores of the groups: Engineering, Business Adminis-

TABLE I—MEANS AND VARIANCES OF THE GROUPS

Major / Scale	Engineering (N = 52)	Physical Education (N = 42)	Education (N = 17)	Business Administration (N = 14)	Between Variance	Within Variance	F	P
Outdoor	49.93	50.93	43.24	31.57	1551.69	241.27	6.43	0.001
Mechanical	53.71	42.24	43.24	35.07	1534.51	104.67	14.66	0.001
Computational	31.75	22.31	22.24	30.79	893.32	62.79	14.23	0.001
Scientific	49.89	38.43	37.65	28.79	2157.44	117.77	18.32	0.001
Persuasive	33.70	39.83	39.29	52.57	1342.56	152.70	8.79	0.001
Artistic	25.12	21.95	30.59	20.29	102.68	53.50	1.92	n.s.
Literary	14.19	14.93	18.65	21.57	256.59	66.04	3.89	0.05
Musical	10.85	10.26	13.29	11.79	105.91	40.09	2.64	n.s.
Social Service	36.29	53.17	44.76	40.29	2266.79	148.12	15.30	0.001
Clerical	42.52	41.64	42.65	57.43	1164.78	175.96	6.62	0.001

tration, Education and Physical Education. Table I illustrates these results.

As shown by Table I, neither the Musical nor the Artistic score distinguished among the four groups of students. The Literary scale distinguished slightly (five per cent l.c.). In order to study these relationships further, the means obtained on each Kuder scale by each academic group were compared. Significant differences among the groups were noted. Table II illustrates the various scales which distinguished each academic group.

DISCUSSION

Although the original hypothesis regarding the degree of similarity between students who demonstrated consistency in their occupational choices and the occupational profiles failed to gain support, certain questions regarding the use of occupational profiles were raised. The occupational profiles have been regarded as evidence of the validity of the Kuder Preference Record. In the long run, these profiles may prove of much greater value in counseling than those based upon *a priori* consideration. Some counselors already employ them in vocational counseling. In so doing they may attempt to match, by inspection, a student's interest profile with that of the occupational group. At another time they may be content to match the significantly high scales. In both methods, the question arises regarding how great a degree of deviation be-

TABLE II—KUDER SCALES DIFFERENTIATING THE GROUPS AND t VALUES

High \ Low	Engineering	t	Physical Education	t	Education	t	Business Administration	t
Engineering			Computational	5.41	Computational	4.20	Outdoor	3.92
			Mechanical	5.76	Mechanical	4.92	Mechanical	5.94
			Scientific	5.10	Scientific	4.04	Scientific	6.45
Physical Education	Persuasive	2.39					Outdoor	4.04
	Social Service	6.83			Social Service	2.40	Social Service	3.43
							Mechanical	2.27
							Scientific	4.67
Education	Social Service	2.52					Outdoor	2.06
							Mechanical	2.21
							Scientific	2.26
Business Administration	Persuasive	5.07	Computational	3.47	Computational	2.99		
	Literary	3.00	Persuasive	3.34	Persuasive	2.98		
	Clerical	3.74	Literary	2.65				
			Clerical	3.86	Clerical	3.30		

$t = 2.00$, $p = 0.05$, $t = 2.60$, $p = 0.01$.

tween the student's profile and the occupational profile should exist before a proposed vocation be excluded from consideration. Another bothersome question which may arise is whether an excessively high interest score on one scale can compensate for a score which is lower than the norm. The reverse of this question is also equally bothersome. One problem which arises regarding the use of these scales is that they are reported in terms of means and no indication of their standard deviation is included in the Kuder Manual. Despite such inherent weaknesses, a sufficient number of vocational counselors feel uneasy when they interpret

Kuder scales. This has led to considerable interest in the development of some sort of statistical or quantitative method of comparing an individual to a norm occupational profile.

This study tends to demonstrate the inapplicability of profile correlation techniques for use with the Kuder. It would seem that a more productive approach to the problem of profile interpretation might lie in the coding of the profiles in a manner similar to that used in the MMPI. This procedure has been suggested by Wiener (*7*) and elaborated upon by Frandsen (*4*).

SUMMARY AND CONCLUSIONS

(1) The Kuder Preference Record was administered to a group of male college students who also listed their choices of occupation in order of preference.

(2) Three trained counselors independently sorted the stated choices into two groups; those which demonstrated a consistent pattern of interest and those which revealed an inconsistent pattern. The groups were next divided according to the student's course of study.

(3) The Kuder profile of each student was compared with the most suitable occupational profile in the Kuder manual. The Du Mas coefficient technique for determining profile similarity was used for this purpose. No significant differences were observed between the occupational profiles and either consistent or inconsistent groups.

(4) As differences in the Kuder scores of the different academic groups were observed, an analysis of variance was calculated. These results revealed that only a few scales were important in distinguishing each group of majors from the others.

(5) Students enrolled in Engineering were distinguished from those enrolled in the other curricula by significantly higher scores on the Mechanical and Scientific scales. This group also attained significantly higher scores on the Computational scale than did either the Education or Physical Education group.

(6) Physical Education students were distinguished primarily by attaining higher Social Service scores than either the Business Administration or Engineering group. The Physical Education majors also achieved slightly higher Social Service scores than did other Education majors.

(7) Students enrolled in Business Administration attained

higher scores on the Persuasive and Clerical scales than did the other three groups. They also scored significantly higher on the Computational scale than did either the Physical Education or Education group. As a group the business students also tended to score higher on the Literary scale, although this trend is not highly significant.

(8) The group which consisted of students majoring in Education was not distinguished by any consistent pattern. As the students in this group were preparing to teach a wide range of subjects, it seems likely that they represented a more heterogeneous group than did the others.

(9) These findings indicate that the use of statistical methods for determining the coefficient of profile similarity between an individual's interest profile and an occupational profile may result in misinterpretation. It is suggested that coding the various interest scales in the manner of the MMPI may provide an answer. Such a procedure would enable a weighting of those scales which are pertinent and the elimination of those which are not pertinent to a given occupation.

REFERENCES

1) E. S. Bordin, "A Theory of Vocational Interests as a Dynamic Phenomena," *Educational and Psychological Measurement*, 3: 49–65, 1943.

2) H. D. Carter, "The Development of Vocational Attitudes," *Journal of Consulting Psychology*, 4: 185–191, 1940.

3) F. M. Du Mas, "The Coefficient of Profile Similarity," *Journal of Clinical Psychology*, 5: 123–131, 1949.

4) A. N. Frandsen, "A Note on Wiener's Coding of Kuder Preference Record Profiles," *Educational and Psychological Measurement*, 12: 137–139, 1952.

5) G. F. Kuder, *Examiner Manual for the Kuder Preference Record*, Chicago, Science Research Associates, 1951.

6) E. K. Strong, Jr., *Vocational Interests of Men and Women*. Stanford, California, Stanford University Press, 1943.

7) D. N. Wiener, "Empirical Occupational Groupings of Kuder Preference Record Profiles," *Educational and Psychological Measurement*, 11: 273–279, 1951.

BOOK REVIEWS

LEOPOLD BELLAK. *The Thematic Apperception Test and The Children's Apperception Test in Clinical Use.* New York: Grune and Stratton, 1954, pp. 282.

This is a book on the TAT and CAT "meant primarily to be of practical use to the student and practitioner of clinical psychology and psychiatry." The volume takes up in eleven chapters and an appendix the following subjects: theoretical foundations for projective testing; clinical use of the TAT; interpretation of the TAT; the TAT pictures as stimuli; clinical and other special problems; character and defenses in the TAT; use of the TAT in psychotherapy; future developments of TAT; the CAT, nature and purpose of the test, history and theory; interpretation of the CAT; a supplement to the CAT, called CAT-S, its purpose, administration, etc.; publications and work in progress on the CAT; tabulations and schemes for CAT and for CAT-S. There is an excellent index, pp. 275–282.

The greatest value of the book probably lies in the detailed discussions of clinical uses, interpretations and special problems in relation to a large number of cases. This value is found further in the fact that CAT is also presented in similar fashion with some comparisons between TAT and CAT.

Many cases are reported, analyses appear in facsimile printed forms of the Analysis Sheet, and details appear including the descriptive theme, interpretative theme and diagnostic level. Clinical notes follow.

The entire work is the result of thorough preparation and training and long experience in one of the best clinical settings. One cannot go through the details of this volume without becoming well acquainted with much that is being done with the tests. These details stand on their own feet.

It is a question as to whether some of the broad generalizations can be accepted. The experimentally oriented psychologist may look with some wonder at the statement that "The main efforts in American Psychology are directed toward making projective methods the tools of nomothetic science, ... " Whether or not they can agree with this, it does not detract from the mass of clinical material that is of great value for the student of projective testing.

This is a valuable contribution in a special field and in the efforts at progress in the stream of American psychology.

A. S. Edwards

The University of Georgia

Leonard Kornberg. *A Class for Disturbed Children: A Case Study and Its Meaning for Education.* New York: Teachers College, Columbia University, 1955, pp. 157.

The work of special classes for children who do not fit into the regular school work finds contribution in this study, which is advertised as a book for all teachers. It is "a case study of a class" in a residential institution. The data were gathered during a five-month period.

The fifteen children, all boys, had failed in other schools. "I was met in my classroom with brooding fear and hate." This was "a residential school for disturbed children, all away from home, all under psychiatric care." The author states that out of his experience came "the insight for a basic thing, a truth to renew and care for in teaching." What was done he believes was a "renewal of the oldest form of teaching—teaching by dialogue" (p. 2). This was a dialogue of "persons," not of "words."

Chapter 1 is Teaching as Dialogue, and Chapter 2, The School's Community. The major part of the book lies in the details of Chapter 3, entitled The Classroom History. The three later chapters explain Methodology and Problems, Some Practical Implications, and Toward General Principles.

The details of the classroom history indicate how clarity came in the method of teaching; these are the bases for the theory worked out in the later chapters. It seems to be a rediscovery of the best in teaching that has been known and used by the best teachers of many ages. Many readers may discover or rediscover it with the author. It may appear to some that this best kind of teaching does not need the complexity of theory that it now receives but it is heartening to know that it is being rediscovered—perhaps that it has not really been lost.

The report presents what is considered as a good method, worked out in actual practice, and given in detail for those who are interested in the best kind of teaching. It is well worth the consideration of those who wish to think through the problems of teaching and

to learn about one of the ways in which a good deal of success has been reached.

It is a pity that so good a book has been marred by the comment about how "poor" our schools have been "for our day" (p. 134). One is reminded of the common mass criticism which all too frequently falls upon our schools. This error (of false aggregation, according to the logicians) does not affect the main presentation. But one who had visited some of our best schools might as reasonably say, "How excellent have been our schools in our day." Some of our schools are poor, some are mediocre, but also some are excellent.

A brief bibliography of one page indicates books that are recommended by the author. There is no index.

A. S. Edwards

The University of Georgia

PUBLICATIONS RECEIVED

EDWARD E. CURETON. *Proceedings 1954 Invitational Conference on Testing Problems.* Princeton, N. J.: Educational Testing Service, pp. 135 (paper).

WINIFRED DE KOK. *You and Your Child.* New York: Philosophical Library Inc., 1955, pp. 147. $3.75.

DIRECCAO DE VITOR FONTES. *A Crianca Portuguesa.* Lisboa: Este Livro Foi Composto E Impresso Na Grafica Santelmo, Rua De S. Bernardo, 84 (paper).

PAUL L. DRESSEL. *Evaluation in General Education.* Dubuque, Iowa: Wm. C. Brown Co., 1954, pp. 333 (paper). $4.00.

PAUL L. DRESSEL AND LEWIS B. MAYHEW. *Critical Thinking in Social Science.* Dubuque, Iowa: Wm. C. Brown Company, 1954, pp. 36 (paper). $1.25.

PAUL L. DRESSEL AND LEWIS B. MAYHEW. *A Handbook for Theme Analysis.* Dubuque, Iowa: Wm. C. Brown Company, 1954, pp. 78 (paper). $1.50.

PAUL L. DRESSEL AND LEWIS B. MAYHEW. *Science Reasoning and Understanding.* Dubuque, Iowa: Wm. C. Brown Company, 1954, pp. 223 (paper). $3.50.

WALTER N. DUROST AND STELLA S. CENTER. *Durost-Center Word Mastery Test.* Yonkers-on-Hudson, N. Y.: World Book Company (paper).

HENRY S. DYER AND RICHARD G. KING. *College Board Scores.* College Entrance Examination Board, 1955, pp. 192 (paper). $1.50.

RUTH E. ECKERT AND ROBERT J. KELLER, Editors. *A University Looks at Its Program: The Report of the University of Minnesota Bureau of Institutional Research, 1942-52.* Minneapolis: The University of Minnesota Press, 1954, pp. 223. $4.00.

Education for Personal and Family Living: A Working Guide for Colleges. A Project of The American Social Hygiene Association. pp. 145 (paper). $1.00.

ALLEN L. EDWARDS. *Edwards Personal Preference Schedule.* New York: The Psychological Corporation, 1954, pp. 18 (paper).

ALLEN L. EDWARDS. *Statistical Methods for the Behavioral Sciences.* New York: Rinehart & Co., Inc., 1954, pp. 542. $6.50.

HENRY EHLERS, EDITOR. *Crucial Issues in Education: An Anthology.* New York: Henry Holt & Co., 1955, pp. 277 (paper). $2.25.